VECTORES PROPIOS GENERALIZADOS Y FORMA CANÓNICA DE JORDAN

Paola Andrea Castro Martínez

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA
MONTERÍA
2020

VECTORES PROPIOS GENERALIZADOS Y FORMA CANÓNICA DE JORDAN

Paola Andrea Castro Martínez

Trabajo presentado como requisito parcial para optar al título de

Matemático

Asesor:

M. Sc. Ricardo Miguel Guzmán Navarro

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA
MONTERÍA
2020

UNIVERSIDAD DE CÓRDOBA
FACULTAD DE CIENCIAS BÁSICAS
DEPARTAMENTO DE MATEMÁTICAS Y ESTADÍSTICA

Los jurados abajo firmantes certifican que han leído y que aprueban el trabajo de grado titulado: **VECTORES PROPIOS GENERALIZADOS Y FORMA CANÓNICA DE JORDAN**, el cual es presentado por la estudiante **Paola Andrea Castro Martínez**.

Fecha: Mayo de 2020

Ricardo Guzman

Asesor: ______________________________
M. Sc. Ricardo Miguel Guzmán Navarro

Jurado: ______________________________
Dr. Carlos Alberto Banquet Brango

Jurado: ______________________________
Dr. Jerson Manuel Borja Soto

A Verónica Rodriguez (in memoriam)
Y a Ofelia Zabaleta

Resumen

En el presente trabajo hacemos uso de los vectores propios generalizados para deducir la Forma canónica de Jordan de una matriz cuadrada, y mostramos algunas aplicaciones de esta. Para ello, demostramos que $\mathbb{C}^n$ se puede descomponer como suma directa de subespacios propios generalizados haciendo uso de los polinomios anuladores y el polinomio minimal. Probamos además que cada subespacio propio generalizado se puede descomponer como suma directa de subespacios cíclicos de Jordan. Finalmente, usamos las descomposiciones anteriormente mencionadas para demostrar el Teorema de Jordan.

Palabras claves: Polinomios anuladores, Vectores propios generalizados, Forma canónica de Jordan.

Abstract

In this work, generalized eigenvalues are used to study the Jordan normal form and some applications of this are shown. We prove that $\mathbb{C}^n$ can be decomposed as a direct sum of generalized proper subspaces by using annihilating polynomials and the minimal polynomial. We also prove that each generalized eigensubspace can be decomposed as a direct sum of Jordan cyclic subspaces. Finally, the Jordan Theorem is proved by using the above decompositions.

Keywords: Annihilator polynomials, Generalized eigenvalues, Jordan normal form.

Agradecimientos

Quiero hacer uso de estas líneas para agradecer primero a Dios por ser mi guía espiritual a lo largo de toda mi carrera, a mis padres Jaime Castro y Martha Martínez y a mis hermanas Verónica Castro y Camila Castro, pues ellos han sido mi soporte durante mis estudios y me han brindado ese amor familiar que solo los padres y hermanas comprenden; a mi sobrino Juan Sebastián, pues su llegada a este mundo generó en mí una nueva motivación para alcanzar este sueño, a mis compañeros de estudio, en especial a mi novio y colega Juan David Barajas, por su apoyo moral y amor incondicional a lo largo de estos años, a la Universidad de Córdoba pues me brindó una formación integral y de calidad durante mis estudios de pregrado, a mi asesor, Ricardo Miguel Guzmán Navarro por todos sus valiosos consejos, comentarios y sugerencias que me ayudaron enormemente al desarrollo de este trabajo, a los profesores del Departamento de Matemáticas y Estadística, pues me enseñaron la belleza de las Matemáticas, a mi jurado Carlos Banquet Brango por sus fundamentales correcciones que ayudaron a finiquitar este trabajo y en especial, quiero agradecer al profesor Jerson Borja Soto, pues además de ser mi jurado de tesis, mi mentor y guía a lo largo de mis estudios, es mi mejor amigo.

Montería, Colombia — Paola Andrea Castro Martínez

Mayo de 2020

Índice general

Introducción

Los valores propios son a menudo introducidos en el contexto del álgebra lineal o teoría matricial. Históricamente, sin embargo, ellos surgen en el estudio de las formas cuadráticas y ecuaciones diferenciales.

En el siglo XVIII, Euler estudió el movimiento rotacional de un cuerpo rígido y descubrió la importancia de los ejes principales. Lagrange generalizó que los ejes principales son en realidad los vectores propios de la matriz de inercia. En el siglo XIX, Cauchy vió como su trabajo podría ser usado para clasificar las superficies cuádricas, y generalizarlo a dimensiones arbitrarias. Cauchy también acuñó el término de "*racine caractéristique*" (raíz característica) que ahora es conocido como valor propio; su término sobrevive en la ecuación característica [6]. A comienzos del siglo XX, Hilbert estudió los valores propios de operadores integrales viendo los operadores como matrices infinitas. Él fue el primero en usar la palabra alemana "*eigen*", la cual significa "*propio*", para denotar los valores y vectores propios en 1904, la cual se sigue usando en la actualidad [1]. Luego, a mediados del siglo XX, este concepto fue generalizado y se obtuvieron relaciones, hasta ese entonces desconocidas, con la forma canónica de Jordan.

Esta generalización, sin embargo, tiene muchos precedentes que ameritan discusión. Como se expresó anteriormente, la generalización de los valores propios está fundamentada en la forma canónica de Jordan la cual tiene un origen bastante particular. En el año 1874 el matemático e ingeniero francés Carmille Jordan (1838-1922) y el matemático alemán Leopold Kronecker (1823-1891) tuvieron una controversia generada por la ambición que tenía Jordan de reorganizar la Teoría de las formas bilineales mediante lo que él denominó noción "simple" de *forma canónica* [2]. La cuestión se plantea como sigue:

Dados dos polinomios bilineales,

$$P = \sum_{i,j=1}^{n} A_{ij} x_i y_j, \quad \text{y} \quad Q = \sum_{s,t=1}^{n} B_{st} x_s y_t,$$

uno se plantea la posibilidad de reducir ambos polinomios P y Q simultáneamente a su forma canónica.

La cuestión fue entonces qué método se debía emplear para organizar la teoría de formas bilineales. Por un lado, Jordan argumentó que el método principal de la teoría debía ser la reducción algebraica de las formas en sus formas canónicas más simples, método que él había desarrollado en el año 1870 en su trabajo *Théorie des substitutions* [3]. Por otra parte, Kronecker objetó que la teoría pertenecía a la aritmética de las formas según la tradición de Gauss, y que, por tanto, el problema principal debía ser caracterizar las clases de equivalencia de las formas mediante el cálculo de invariantes.

Como podemos evidenciar, el enfoque de Jordan fue el más aceptado por la comunidad matemática debido a que era más simple y elegante. Este es uno de los múltiples ejemplos que existen sobre el rechazo de las ideas excéntricas de Kronecker por los matemáticos de su época. Con la forma canónica de Jordan se exploraron nuevos campos de investigación y empezaron a encontrarse relaciones en otras áreas de las matemáticas tales como el Análisis Numérico y los Sistemas Dinámicos.

En este trabajo se utilizará la teoría de vectores propios generalizados para hallar la Forma canónica de Jordan. En primer lugar, en el Capítulo 1, se presentan los conceptos y teoremas básicos que serán necesarios para entender esta monografía. En el Capítulo 2, se analiza detalladamente el conjunto de polinomios que anulan a una matriz dada $\mathbf{A}$ y mostramos el teorema de Cayley-Hamilton que establece que toda matriz satisface su polinomio característico y, por lo tanto, el conjunto antes mencionado es no vacío. Más adelante, en el Capítulo 3, se presenta la teoría de subespacios propios generalizados y encontramos una descomposición de $\mathbb{C}^n$ como suma directa de subespacios propios generalizados, los cuales, a su vez, se descomponen en suma directa de subespacios cíclicos de Jordan. Finalmente, en el Capítulo 4, se prueba el

Teorema de Jordan, el cual es el principal objetivo de este trabajo, y se proporcionan algunas aplicaciones de este resultado.

Capítulo 1

Preliminares

En este capítulo mostramos una serie de definiciones y se enunciarán teoremas que serán usados más adelante. Daremos la demostración de algunos resultados. Los temas que trataremos en este capítulo son matrices, espacios vectoriales, polinomios y valores y vectores propios, los cuales son estudiados en los cursos de Álgebra Lineal y Álgebra Abstracta. Se ha decidido omitir la demostración de algunos resultados para que este trabajo no se haga muy extenso, sin embargo, estas pueden consultarse en [4, 5, 7, 8, 9, 11].

Todos los resultados de este trabajo se muestran sobre el campo de los números complejos, pero en realidad son válidos para cualquier campo algebraicamente cerrado.

1.1. Números Complejos

El símbolo $\mathbb{C}$ representa al conjunto de los **números complejos**. Todo número complejo se representa en la forma $a + ib$ donde $a, b \in \mathbb{R}$. La suma y multiplicación de números complejos se definen por las ecuaciones siguientes:

$$(a + ib) + (c + id) = (a + c) + i(b + d), \quad (a + ib)(c + id) = (ac - bd) + i(ad + bc),$$

donde $a, b, c, d \in \mathbb{R}$. El conjunto $\mathbb{C}$ con esta suma y multiplicación forma un campo. Esto es, se satisfacen las propiedades siguientes:

1. *Conmutatividad:* $\alpha + \beta = \beta + \alpha$ y $\alpha\beta = \beta\alpha$, para todo $\alpha, \beta \in \mathbb{C}$;

2. *Asociatividad:* $(\alpha+\beta)+\lambda = \beta+(\alpha+\lambda)$ y $(\alpha\beta)\lambda = \alpha(\beta\lambda)$, para todo $\alpha, \beta, \lambda \in \mathbb{C}$;

3. *Identidades:* $\alpha + 0 = \alpha$ y $\alpha 1 = \alpha$ para todo $\alpha \in \mathbb{C}$;

4. *Inverso aditivo:* para cada $\alpha \in \mathbb{C}$, existe un único $\beta \in \mathbb{C}$ tal que $\alpha + \beta = 0$. El elemento β es llamado el **inverso aditivo de** α y se denota por $-\alpha$;

5. *Inverso multiplicativo:* para cada $\alpha \in \mathbb{C} - \{0\}$, existe un único $\beta \in \mathbb{C}$ tal que $\alpha\beta = 1$. Este elemento β es llamado el **inverso multiplicativo de** α y se denota por α^{-1};

6. *Propiedad distributiva:* $\alpha(\beta + \lambda) = \alpha\beta + \alpha\lambda$ para todo $\alpha, \beta, \lambda \in \mathbb{C}$.

La **resta y división** de números complejos $\alpha, \beta \in \mathbb{C}$ se definen, respectivamente, mediante las ecuaciones

$$\beta - \alpha = \beta + (-\alpha), \quad \beta/\alpha = \beta\alpha^{-1},$$

donde se supone $\alpha \neq 0$ en la última ecuación.

1.2. Matrices

Sean m y n enteros positivos. Una **matriz de tamaño** $m \times n$ es un arreglo rectangular de elementos en $\mathbb{C}$ de la forma

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1n} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \mathbf{A}_{m2} & \cdots & \mathbf{A}_{mn} \end{bmatrix}$$

Los elementos $\mathbf{A}_{ij}$ son llamados las **componentes o entradas** de $\mathbf{A}$. Para $1 \leq i \leq m$, la $i-$**ésima fila** de $\mathbf{A}$ es

$$\mathbf{A}_{i*} = \begin{bmatrix} \mathbf{A}_{i1} & \mathbf{A}_{i2} & \cdots & \mathbf{A}_{in} \end{bmatrix},$$

y para $1 \leq j \leq n$, la $j-$**ésima columna** de $\mathbf{A}$ es

$$\mathbf{A}_{*j} = \begin{bmatrix} \mathbf{A}_{1j} \\ \mathbf{A}_{2j} \\ \vdots \\ \mathbf{A}_{mj} \end{bmatrix}.$$

El conjunto formado por todas las matrices de tamaño $m \times n$ se denotará por el símbolo $\mathbb{C}^{m\times n}$.

Definición 1.2.1. Sean $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$. Se define la **suma de A y B** como la matriz $\mathbf{A} + \mathbf{B} \in \mathbb{C}^{m\times n}$ tal que

$$(\mathbf{A} + \mathbf{B})_{ij} = \mathbf{A}_{ij} + \mathbf{B}_{ij}.$$

La suma de matrices posee las propiedades que mencionamos a continuación:

1. La suma de matrices es conmutativa y asociativa.

2. La matriz $m \times n$ que tiene todas sus componentes iguales a cero se llama la **matriz nula** y se representa por el símbolo 0. La matriz 0 tiene la propiedad de que para toda $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{A} + 0 = \mathbf{A}$.

3. Si $\mathbf{A} \in \mathbb{C}^{m\times n}$, la **matriz opuesta** de $\mathbf{A}$ es la matriz de tamaño $m\times n$, denotada por $-\mathbf{A}$, tal que $(-\mathbf{A})_{ij} = -\mathbf{A}_{ij}$. La matriz opuesta de $\mathbf{A}$ satisface que $\mathbf{A} + (-\mathbf{A}) = 0$.

Definición 1.2.2. Si $\mathbf{A} \in \mathbb{C}^{m\times n}$ y $\alpha \in \mathbb{C}$, se define la **multiplicación** de α por $\mathbf{A}$ como la matriz $\alpha\mathbf{A} \in \mathbb{C}^{m\times n}$ tal que

$$(\alpha\mathbf{A})_{ij} = \alpha\mathbf{A}_{ij}.$$

Esta operación posee las siguientes propiedades, donde $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$ y $\alpha, \beta \in \mathbb{C}$:

1. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$;

2. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$;

3. $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$;

4. $1\mathbf{A} = \mathbf{A}$;

Definición 1.2.3. Para matrices $\mathbf{A} \in \mathbb{C}^{m\times n}$ y $\mathbf{B} \in \mathbb{C}^{n\times p}$, se define la **multiplicación de A y B** como la matriz $\mathbf{AB} \in \mathbb{C}^{m\times p}$ que satisface

$$(\mathbf{AB})_{ij} = \sum_{k=1}^{n} \mathbf{A}_{ik}\mathbf{B}_{kj}.$$

Con la multiplicación así definida se tienen las siguientes propiedades, donde $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times p}$, $\mathbf{C} \in \mathbb{C}^{p\times q}$:

1. $(\mathbf{AB})_{i*} = \mathbf{A}_{i*}\mathbf{B}$;
2. $(\mathbf{AB})_{i*} = \mathbf{A}_{i1}\mathbf{B}_{1*} + \cdots + \mathbf{A}_{in}\mathbf{B}_{n*}$;
3. $(\mathbf{AB})_{*j} = \mathbf{AB}_{*j}$;
4. $(\mathbf{AB})_{*j} = \mathbf{B}_{1j}\mathbf{A}_{*1} + \cdots + \mathbf{B}_{nj}\mathbf{A}_{*n}$;
5. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

Además, para $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m\times n}$, y $\mathbf{C} \in \mathbb{C}^{n\times p}$ tenemos

1. $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$;
2. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.

La **matriz identidad** de $\mathbb{C}^{n\times n}$, denotada por el símbolo $\mathbf{I}_n$, está definida por

$$(\mathbf{I}_n)_{ij} = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{si } i \neq j. \end{cases}$$

La $i-$ésima columna de $\mathbf{I}_n$ será denotada por el símbolo $\mathbf{e}_i$. Para una matriz $\mathbf{A} \in \mathbb{C}^{m\times n}$ se tienen las igualdades siguientes:

$$\mathbf{AI}_n = \mathbf{A} \text{ e } \mathbf{I}_m\mathbf{A} = \mathbf{A}$$

Definición 1.2.4. Una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es **invertible** si existe $\mathbf{B} \in \mathbb{C}^{n\times n}$ tal que

$$\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}.$$

Cuando $\mathbf{A} \in \mathbb{C}^{n\times n}$ es invertible, entonces existe una única matriz $\mathbf{B} \in \mathbb{C}^{n\times n}$ que satisface $\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}$. Esta matriz se denota por $\mathbf{A}^{-1}$ y se llama **la inversa de A**. Si $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ son matrices invertibles, entonces $\mathbf{A}^{-1}$ y $\mathbf{AB}$ son invertibles y

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \quad (\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Definición 1.2.5. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Denotaremos por $\mathbf{A}^T \in \mathbb{C}^{n\times n}$ a la matriz cuya entrada ij es $\mathbf{A}_{ji}$, para todo i, j, y la llamaremos **matriz transpuesta** de $\mathbf{A}$.

Teorema 1.2.1. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$ invertible. Entonces $\mathbf{A}^T$ es invertible y $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.

Las potencias $\mathbf{A}^k$ de una matriz no nula $\mathbf{A} \in \mathbb{C}^{n\times n}$, donde k es un entero positivo, se definen inductivamente como sigue:

1. $\mathbf{A}^1 = \mathbf{A}$ y

2. $\mathbf{A}^{k+1} = \mathbf{A}^k\mathbf{A}$.

Por conveniencia se define $\mathbf{A}^0 = I_n$. Se satisfacen las propiedades de los exponentes siguientes, donde $\mathbf{A} \in \mathbb{C}^{n\times n}$ y k, l son enteros positivos:

1. $\mathbf{A}^k\mathbf{A}^l = \mathbf{A}^{k+l}$;

2. $(\mathbf{A}^k)^l = \mathbf{A}^{kl}$.

Una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es llamada **triangular superior** si $\mathbf{A}_{ij} = 0$ siempre que $i > j$. Una **matriz diagonal** es una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ que satisface que $\mathbf{A}_{ij} = 0$ siempre que $i \neq j$. Para una matriz $\mathbf{A} \in \mathbb{C}^{m\times n}$, su **transpuesta hermitiana**, es la matriz $\mathbf{A}^H \in \mathbb{C}^{n\times m}$ dada por:

$$(\mathbf{A}^H)_{ij} = \overline{\mathbf{A}_{ji}}.$$

Definición 1.2.6. Una matriz $\mathbf{U} \in \mathbb{C}^{n\times n}$ es llamada **unitaria** si $\mathbf{U}^H = \mathbf{U}^{-1}$.

Definición 1.2.7. Sean $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$. Decimos que $\mathbf{A}$ es **semejante** a $\mathbf{B}$ si existe $\mathbf{P} \in \mathbb{C}^{n\times n}$ invertible tal que

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{B}.$$

La semejanza de matrices es una relación de equivalencia en $\mathbb{C}^{n\times n}$.

Sean $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ matrices semejantes y $k \in \mathbb{Z}^+$. Entonces $\mathbf{A}^k$ y $\mathbf{B}^k$ son semejantes. Además, si $\lambda \in \mathbb{C}$, entonces $\mathbf{A} - \lambda\mathbf{I}_n$ y $\mathbf{B} - \lambda\mathbf{I}_n$ son semejantes.

Definición 1.2.8. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Decimos que $\mathbf{A}$ es **diagonalizable** si es semejante a una matriz diagonal.

1.3. El espacio vectorial $\mathbb{C}^n$

El símbolo $\mathbb{C}^n$ denotará al conjunto de matrices $\mathbb{C}^{n\times 1}$ y sus elementos serán llamados **vectores**.

Definición 1.3.1. Un subconjunto no vacío W de $\mathbb{C}^n$ es un **subespacio** de $\mathbb{C}^n$ si para todo par de vectores $\mathbf{w}_1, \mathbf{w}_2 \in W$ y todo $c \in \mathbb{C}$, el vector $c\mathbf{w}_1 + \mathbf{w}_2$ está en W.

Definición 1.3.2. Sean $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}^n$. Una **combinación lineal** de los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ es una expresión de la forma

$$c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k$$

donde $c_1, \ldots, c_k \in \mathbb{C}$.

Lema 1.3.1. Sean $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}^n$. El conjunto formado por todas las combinaciones lineales de los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ es un subespacio de $\mathbb{C}^n$.

Demostración. Véase [8, página 37] □

Definición 1.3.3. Sean $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ vectores en $\mathbb{C}^n$. El subespacio de todas las combinaciones lineales de $\mathbf{x}_1, \cdots, \mathbf{x}_k$ es llamado el **generado por los vectores** $\mathbf{x}_1, \ldots, \mathbf{x}_k$, y lo denotamos por el símbolo $\operatorname{gen}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$.

Definición 1.3.4. Un subconjunto $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ de $\mathbb{C}^n$ es **linealmente independiente** si la única manera de expresar al vector 0 de $\mathbb{C}^n$ como combinación lineal de los vectores $\mathbf{x}_1, \ldots, \mathbf{x}_k$ es que todos los escalares sean iguales a cero.

Definición 1.3.5. Sea W un subespacio de $\mathbb{C}^n$. Una **base** de W es un conjunto $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ de vectores de $\mathbb{C}^n$ tales que:

1. $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ es linealmente independiente;

2. $\text{gen}\,\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = W$.

Teorema 1.3.2. Sea W un subespacio de $\mathbb{C}^n$. Dos bases cualesquiera de W tienen el mismo número (finito) de vectores. Ese número común es llamado la **dimensión** de W, y lo denotaremos por el símbolo $\dim(W)$.

Demostración. Véase [8, página 44]. $\square$

Lema 1.3.3. La dimensión del espacio $\mathbb{C}^n$ es n.

Demostración. Las columnas de $\mathbf{I}_n$ forman una base de $\mathbb{C}^n$, luego por el Teorema 1.3.2, $\dim \mathbb{C}^n = n$. $\square$

Lema 1.3.4. Para cualquier subespacio W de $\mathbb{C}^n$ se cumple que $\dim(W) \leq n$.

Lema 1.3.5. Sean W_1, W_2 subespacios de $\mathbb{C}^n$ tales que $W_1 \subsetneqq W_2$. Entonces

$$\dim(W_1) < \dim(W_2).$$

Demostración. Véase [8, página 46]. $\square$

Definición 1.3.6. Sean $W_1, \ldots, W_k$ subespacios de $\mathbb{C}^n$. Entonces el conjunto

$$W = W_1 + \cdots + W_k = \{\mathbf{w}_1 + \cdots + \mathbf{w}_k : \mathbf{w}_1 \in W_1, \ldots, \mathbf{w}_k \in W_k\}$$

es llamado la **suma de los subespacios** $W_1, \ldots, W_k$.

En la Definición 1.3.6, W es un subespacio de $\mathbb{C}^n$ que contiene a cada uno de los W_i (véase [8, página 37]).

Teorema 1.3.6. Sean $W_1, \ldots, W_k$ subespacios de $\mathbb{C}^n$ y $W = W_1 + \cdots + W_k$. Las siguientes afirmaciones son equivalentes.

1. $\mathbf{w}_1 + \cdots + \mathbf{w}_k = 0$, donde $\mathbf{w}_i \in W_i$ para todo $i \in \{1, 2, \ldots, k\}$, implica que $\mathbf{w}_i = 0$, $i = 1, 2, \ldots, k$.

2. Para todo $j \in \{1, \ldots, k\}$, $W_j \cap \sum\limits_{\substack{i=1 \\ i \neq j}}^{k} W_i = \{0\}$.

3. Si β_i es una base de W_i para todo $i \in \{1, \ldots, k\}$. Entonces

$$\beta = \beta_1 \cup \cdots \cup \beta_k$$

es una base de W.

Demostración. Véase [8, página 208]. □

Definición 1.3.7. Si cualquiera de las condiciones del Teorema 1.3.6 se cumple, se dice que la suma $W = W_1 + \cdots + W_k$ es **directa** o que W es la **suma directa** de $W_1, \ldots, W_k$, y se escribe

$$W = W_1 \oplus \cdots \oplus W_k.$$

En el resto de esta sección trataremos con subespacios que están relacionados con matrices.

Lema 1.3.7. Una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es invertible si y solo si sus columnas forman una base de $\mathbb{C}^n$.

Demostración. Véase [8, página 46] □

Lema 1.3.8. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^n : \mathbf{A}\mathbf{x} = 0\},$$

es un subespacio de $\mathbb{C}^n$.

Demostración. Como $\mathbf{A}0 = 0$, entonces $0 \in \mathcal{N}(\mathbf{A})$ y así $\mathcal{N}(\mathbf{A})$ es no vacío. Ahora, sean $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$ y $c \in \mathbb{C}$. Entonces

$$\mathbf{A}(c\mathbf{x}_1 + \mathbf{x}_2) = c\mathbf{A}\mathbf{x}_1 + \mathbf{A}\mathbf{x}_2 = c0 + 0 = 0.$$

Esto muestra que $c\mathbf{x}_1 + \mathbf{x}_2 \in \mathcal{N}(\mathbf{A})$. Por lo tanto, $\mathcal{N}(\mathbf{A})$ es un subespacio de $\mathbb{C}^n$. □

Definición 1.3.8. El subespacio $\mathcal{N}(\mathbf{A})$ del Lema 1.3.8 es llamado el **espacio nulo** de $\mathbf{A}$.

Lema 1.3.9. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $k \in \mathbb{Z}^+$. Entonces

$$\mathcal{N}(\mathbf{A}^k) \subseteq \mathcal{N}(\mathbf{A}^{k+1}).$$

Demostración. Sea $\mathbf{x} \in \mathcal{N}(\mathbf{A}^k)$. Entonces $\mathbf{A}^k\mathbf{x} = \mathbf{0}$, y por lo tanto $\mathbf{A}\mathbf{A}^k\mathbf{x} = \mathbf{A}\mathbf{0} = \mathbf{0}$, esto es, $\mathbf{A}^{k+1}\mathbf{x} = \mathbf{0}$. Así, $\mathbf{x} \in \mathcal{N}(\mathbf{A}^{k+1})$. Por lo tanto, $\mathcal{N}(\mathbf{A}^k) \subseteq \mathcal{N}(\mathbf{A}^{k+1})$ □

Lema 1.3.10. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Si $\mathbf{A}\mathbf{x} = 0$ para todo $\mathbf{x} \in \mathbb{C}^n$, entonces $\mathbf{A} = 0$.

Demostración. Sean $\mathbf{e}_1, \ldots, \mathbf{e}_n$ las columnas de la matriz $\mathbf{I}_n$. Entonces, para todo $j \in \{1, \ldots, n\}$ tenemos que $\mathbf{A}_{*j} = \mathbf{A}\mathbf{e}_j = 0$, lo que muestra que todas las columnas de $\mathbf{A}$ son nulas y así $\mathbf{A} = 0$. □

Definición 1.3.9. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. El conjunto

$$\mathcal{C}(\mathbf{A}) = \operatorname{gen}\{\mathbf{A}_{*1}, \ldots, \mathbf{A}_{*n}\}$$

es un subespacio de $\mathbb{C}^n$ llamado el **espacio columna** de $\mathbf{A}$.

Definición 1.3.10. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Definimos el **rango** de $\mathbf{A}$ como la dimensión del espacio columna de $\mathbf{A}$ y lo denotamos por $\operatorname{rk}(\mathbf{A})$.

Lema 1.3.11. Las matrices semejantes tienen el mismo rango.

Demostración. Véase [9, página 46]. □

Teorema 1.3.12. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces

$$\dim(\mathcal{N}(\mathbf{A})) + \dim(\mathcal{C}(\mathbf{A})) = n.$$

Demostración. Véase [11, página 199]. □

Lema 1.3.13. Sea $\mathbf{A} \in \mathbb{C}^{n\times n} - \{0\}$ tal que $\mathcal{N}(\mathbf{A}) \neq \{0\}$. Entonces existe un entero positivo p tal que

$$\{0\} \subsetneqq \cdots \subsetneqq \mathcal{N}(\mathbf{A}^{p-1}) \subsetneqq \mathcal{N}(\mathbf{A}^p) = \mathcal{N}(\mathbf{A}^{p+k})$$

para todo $k \in \mathbb{Z}^+$.

Demostración. Por el Lema 1.3.9 $\mathcal{N}(\mathbf{A}^k) \subseteq \mathcal{N}(\mathbf{A}^{k+1})$ para todo $k \in \mathbb{Z}^+$. Así, tenemos la cadena de contenencias

$$\{0\} \subsetneqq \mathcal{N}(\mathbf{A}) \subseteq \mathcal{N}(\mathbf{A}^2) \subseteq \cdots \subseteq \mathcal{N}(\mathbf{A}^k) \subseteq \cdots$$

Ahora, las dimensiones de los subespacios $\mathcal{N}(\mathbf{A}^k)$, $k > 0$, definen una sucesión de enteros no decreciente

$$0 = \dim\{0\} < \dim(\mathcal{N}(\mathbf{A})) \leq \dim(\mathcal{N}(\mathbf{A}^2)) \leq \cdots \leq \dim(\mathcal{N}(\mathbf{A}^k)) \leq \cdots$$

Ahora, tenemos que para todo $k \in \mathbb{Z}^+$, $\mathcal{N}(\mathbf{A}^k)$ es subespacio de $\mathbb{C}^n$, entonces por el Lema 1.3.4, cada número de la sucesión anterior es menor o igual que n, así que debe existir $s \in \mathbb{Z}^+$ tal que $\dim(\mathcal{N}(\mathbf{A}^s)) = \dim(\mathcal{N}(\mathbf{A}^{s+k}))$ para todo $k \in \mathbb{Z}^+$. Escogemos a p como el menor de tales s. Luego se tiene que

$$0 = \dim\{0\} < \dim(\mathcal{N}(\mathbf{A})) < \cdots < \dim(\mathcal{N}(\mathbf{A}^{p-1})) < \dim(\mathcal{N}(\mathbf{A}^p)) = \dim(\mathcal{N}(\mathbf{A}^{p+k}))$$

para todo $k \in \mathbb{Z}^+$. Y esto equivale a tener

$$\{0\} \subsetneqq \cdots \subsetneqq \mathcal{N}(\mathbf{A}^{p-1}) \subsetneqq \mathcal{N}(\mathbf{A}^p) = \mathcal{N}(\mathbf{A}^{p+k})$$

para todo $k \in \mathbb{Z}^+$. $\square$

1.4. Polinomios con coeficientes $\mathbb{C}$

En esta sección hablaremos sobre polinomios con coeficientes en $\mathbb{C}$, pues en este trabajo se trabaja solo con este tipo de polinomios. Muchos de los conceptos aquí mencionados valen para polinomios con coeficientes otros campos en general.

Definición 1.4.1. Un **polinomio con coeficientes en** $\mathbb{C}$, o simplemente un **polinomio**, es una expresión de la forma

$$p(x) = c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0$$

donde $c_n, c_{n-1}, \ldots, c_1, c_0 \in \mathbb{C}$ y x es una variable que solo puede tomar valores en $\mathbb{C}$. Si $\alpha \in \mathbb{C}$, entonces el símbolo $p(\alpha)$ representa al número complejo $c_n\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0$. Así,

$$p(\alpha) = c_n\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0.$$

El conjunto de todos los polinomios con coeficientes en $\mathbb{C}$ lo denotaremos por el símbolo $\mathbb{C}[x]$.

Un polinomio de la forma $p(x) = c$, donde $c \in \mathbb{C}$, es un **polinomio constante**. Los polinomios constantes $p(x) = 0$ y $q(x) = 1$ se denotarán por 0 y 1, respectivamente.

Por conveniencia, en ocasiones se representa un polinomio $p(x)$ con coeficientes en $\mathbb{C}$ en la forma

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots ,$$

donde se entiende que todos los coeficiente a_n son cero, excepto un número finito de ellos. Esto es útil para definir la suma y multiplicación de polinomios.

Definición 1.4.2. Si $p(x) = a_0+a_1x+\cdots+a_nx^n+\cdots$ y $q(x) = b_0+b_1x+\cdots+b_nx^n+\cdots$ están ambos en $\mathbb{C}[x]$, entonces **la suma de** $p(x)$ **y** $q(x)$ es el polinomio con coeficientes $a_n + b_n$, para todo $n \geq 0$.

Definición 1.4.3. Si $p(x) = a_0 + a_1x + \cdots + a_nx^n + \cdots$ y $q(x) = b_0 + b_1x + \cdots + b_nx^n + \cdots$, están ambos en $\mathbb{C}[x]$, entonces $p(x)q(x)$ es definido como el polinomio que tiene coeficientes c_n, $n \geq 0$, dados por la fórmula

$$c_n = a_0b_n + a_1b_{n-1} + a_2b_{n-2} + \cdots + a_nb_0 = \sum_{k=0}^{n} a_kb_{n-k}.$$

Con la suma y multiplicación de polinomios, el conjunto $\mathbb{C}[x]$ satisface todos los axiomas de un campo a excepción del axioma de existencia de inversos multiplicativos.

Se dan a continuación varias definiciones y teoremas relacionados con polinomios que serán usados en el desarrollo del trabajo.

Definición 1.4.4. Sea

$$p(x) = c_nx^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 \in \mathbb{C}\left[x\right],$$

donde los exponentes de x van decreciendo estrictamente de izquierda a derecha y $c_n \neq 0$. El entero n es llamado el **grado** de $p(x)$, y se denota por $\deg(p(x))$.

Si $p(x)$ y $q(x)$ son polinomios no nulos, entonces

$$\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x)).$$

Definición 1.4.5. Un polinomio $p(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 \in \mathbb{C}[x]$ de grado n es llamado **mónico** si $c_n = 1$.

Definición 1.4.6. Sean $p(x), q(x) \in \mathbb{C}[x]$. Decimos que $p(x)$ **divide** a $q(x)$ si existe $s(x) \in \mathbb{C}[x]$ tal que $q(x) = p(x)s(x)$. Si $p(x)$ divide a $q(x)$, entonces $q(x)$ es llamado un **múltiplo** de $p(x)$.

Teorema 1.4.1 (Algoritmo de la División)**.** Sean $p(x), q(x) \in \mathbb{C}[x]$ con $q(x) \neq 0$. Entonces existen únicos $s(x)$ y $r(x)$ en $\mathbb{C}[x]$ tales que

$$p(x) = s(x)q(x) + r(x),$$

donde $r(x) = 0$ o $\deg(r(x)) < \deg(q(x))$.

Demostración. Véase [7, página 139]. □

Definición 1.4.7. Sean $p(x), q(x) \in \mathbb{C}[x]$ polinomios no nulos. Decimos que $p(x)$ y $q(x)$ son **primos relativos** si el mayor grado de cualquier polinomio que divida a ambos es cero. Esto es, los únicos divisores comunes de $p(x)$ y $q(x)$ son polinomios constantes.

El concepto de polinomios primos relativos se extiende a un número finito de polinomios no nulos $p_1(x), p_2(x), \ldots, p_n(x) \in \mathbb{C}[x]$, donde $p_1(x), p_2(x), \ldots, p_n(x)$ son primos relativos si y solo si los únicos divisores comunes de $p_1(x), p_2(x), \ldots, p_n(x)$ son polinomios constantes.

Teorema 1.4.2. Sean $p_1(x), p_2(x), \ldots, p_n(x) \in \mathbb{C}[x]$ polinomios no nulos primos relativos. Entonces existen $s_1(x), s_2(x), \ldots, s_n(x) \in \mathbb{C}[x]$ tales que

$$p_1(x)s_1(x) + p_2(x)s_2(x) + \cdots + p_n(x)s_n(x) = 1.$$

Demostración. Por simplicidad haremos la prueba para $n = 2$, pero la demostración es análoga para n arbitrario. Consideremos los conjuntos

$$S = \{u(x)p_1(x) + t(x)p_2(x) : u(x), t(x) \in \mathbb{C}[x]\} - \{0\}$$

y

$$E = \{\deg(v(x)) : v(x) \in S\}.$$

Nótese que $E \subseteq \mathbb{Z}^+ \cup \{0\}$ y E es no vacío porque $\deg(p(x)) \in E$. Por el principio del buen orden existe el mínimo de E, por lo que existe $d(x) \in S$ tal que $\deg(d(x)) = \min E$. Veamos que $d(x)$ es un divisor común de $p_1(x)$ y $p_2(x)$. En efecto, como $d(x) \in S$, entonces existen $u_1(x), t_1(x) \in \mathbb{C}[x]$ tales que

$$d(x) = u_1(x)p_1(x) + t_1(x)p_2(x).$$

Ahora, por el algoritmo de la división, $d(x)$ divide a $p_1(x)$ ó existen $u_2(x)$ y $r(x)$ únicos en $\mathbb{C}[x]$ tales que

$$p_1(x) = u_2(x)d(x) + r(x), \text{ donde } r(x) \neq 0 \text{ y } \deg(r(x)) < \deg(d(x)).$$

Si suponemos que lo segundo es lo que ocurre, entonces

$$\begin{aligned} r(x) &= p_1(x) - u_2(x)d(x) \\ &= p_1(x) - u_2(x)\left[u_1(x)p_1(x) + t_1(x)p_2(x)\right] \\ &= p_1(x) - u_2(x)u_1(x)p_1(x) - u_2(x)t_1(x)p_2(x) \\ &= \left[1 - u_2(x)u_1(x)\right]p_1(x) + \left[-u_2(x)t_1(x)\right]p_2(x), \end{aligned}$$

así que $r(x) \in S$ y $\deg(r(x)) < \deg(d(x))$, lo que contradice la minimalidad de $\deg(d(x))$. Luego, lo que ocurre es que $d(x)$ divide a $p_1(x)$. De manera similar se demuestra que $d(x)$ divide a $p_2(x)$. Como $p_1(x)$ y $p_2(x)$ son primos relativos, entonces $d(x) = c$, para algún $c \in \mathbb{C} - \{0\}$. Así,

$$\left[c^{-1}u_1(x)\right]p_1(x) + \left[c^{-1}t_1(x)\right]p_2(x) = 1.$$

Haciendo $s(x) = c^{-1}u_1(x)$ y $s_2(x) = c^{-1}t_1(x)$, tenemos que

$$s_1(x)p_1(x) + s_2(x)p_2(x) = 1.$$

□

Definición 1.4.8. Sean $p(x) \in \mathbb{C}[x]$ y $\alpha \in \mathbb{C}$. Decimos que α es una **raíz** de $p(x)$, o que α **satisface** $p(x)$, si $p(\alpha) = 0$.

Teorema 1.4.3 (Teorema Fundamental de Álgebra)**.** Todo polinomio no constante en $\mathbb{C}[x]$ tiene una raíz en $\mathbb{C}$.

Teorema 1.4.4. Sea $p(x) \in \mathbb{C}[x]$. Entonces $\alpha \in \mathbb{C}$ es una raíz de $p(x)$ si y solo si $q(x) = x - \alpha$ divide a $p(x)$.

Definición 1.4.9. Sean $p(x) \in \mathbb{C}[x]$ con $\deg(p(x)) > 0$ y α una raíz de $p(x)$. La **multiplicidad** de α como raíz de $p(x)$ es el número entero positivo m tal que $(x-\alpha)^m$ divide a $p(x)$ y $(x-\alpha)^{m+1}$ no divide a $p(x)$.

Teorema 1.4.5. Sea $p(x) \in \mathbb{C}[x]$ un polinomio mónico de grado $n > 0$. Si $\lambda_1, \dots, \lambda_k$ son las distintas raíces de $p(x)$ con multiplicidades $m_1, \dots, m_k$, respectivamente, entonces

$$p(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}.$$

Además, $m_1 + \cdots + m_k = n$.

Teorema 1.4.6. Sean $p(x), q(x) \in \mathbb{C}[x] - \{0\}$ tales que $p(x)$ divide a $q(x)$. Entonces toda raíz α de $p(x)$ es también raíz de $q(x)$. Además, la multiplicidad de α como raíz de $p(x)$ es menor o igual que la multiplicidad de α como raíz de $q(x)$.

1.5. Valores y vectores propios

Definición 1.5.1. Sean $\mathbf{A} \in \mathbb{C}^{n \times n}$ y $\lambda \in \mathbb{C}$. Decimos que λ es un **valor propio** de $\mathbf{A}$ si existe $\mathbf{x} \in \mathbb{C}^n - \{0\}$ tal que

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x}.$$

En esta situación diremos que $\mathbf{x}$ es un **vector propio** de $\mathbf{A}$ asociado al valor propio λ. El **espectro de A** es el conjunto

$$\mathrm{Spec}(\mathbf{A}) = \{\lambda \in \mathbb{C} : \lambda \text{ es un valor propio de } \mathbf{A}\}.$$

Si $\mathbf{A} \in \mathbb{C}^{n \times n}$ y λ es un valor propio de $\mathbf{A}$, entonces el conjunto

$$E_{\mathbf{A}}(\lambda) = \{\mathbf{x} \in \mathbb{C}^n : \mathbf{A}\mathbf{x} = \lambda \mathbf{x}\}$$

es un subespacio de $\mathbb{C}^n$. Nótese que

$$E_{\mathbf{A}}(\lambda) = \{0\} \cup \{\mathbf{x} \in \mathbb{C}^n : \mathbf{x} \text{ es un vector propio de } \mathbf{A} \text{ asociado a } \lambda\}.$$

El subespacio $E_{\mathbf{A}}(\lambda)$ de $\mathbb{C}^n$ es llamado el **subespacio propio de A** asociado al valor propio λ.

Teorema 1.5.1. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces $\lambda \in \mathbb{C}$ es un valor propio de $\mathbf{A}$ si y solo si $\det(\lambda \mathbf{I}_n - \mathbf{A}) = 0$, donde $\det(\lambda \mathbf{I}_n - \mathbf{A})$ es el determinante de la matriz $\lambda \mathbf{I}_n - \mathbf{A}$

Demostración. Véase [9, página 38]. $\square$

Definición 1.5.2. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. El **polinomio característico** de $\mathbf{A}$, denotado $P_{\mathbf{A}}(x)$ es el polinomio en $\mathbb{C}[x]$ definido por

$$P_{\mathbf{A}}(x) = \det(x\mathbf{I}_n - \mathbf{A}).$$

Definición 1.5.3. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y λ un valor propio de $\mathbf{A}$. La **multiplicidad algebraica** de λ, denotada por $\mathrm{ma}_{\mathbf{A}}(\lambda)$ es la multiplicidad de λ como raíz del polinomio característico $P_{\mathbf{A}}(x)$.

Definición 1.5.4. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y λ un valor propio de $\mathbf{A}$. La **multiplicidad geométrica** de λ, denotada $\mathrm{mg}_{\mathbf{A}}(\lambda)$, se define como la dimensión de $E_{\mathbf{A}}(\lambda)$.

Teorema 1.5.2. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces $P_{\mathbf{A}}(x) = \det(x\mathbf{I}_n - \mathbf{A})$ es un polinomio de grado n.

Demostración. Véase [9, página 38] $\square$

Si $p(x) = a_0 + a_1x + \cdots + a_kx^k$ es un polinomio con coeficientes en $\mathbb{C}$ y $\mathbf{A} \in \mathbb{C}^{n\times n}$, se define la matriz $p(\mathbf{A})$ mediante

$$p(\mathbf{A}) = a_0I_n + a_1\mathbf{A} + \cdots + a_k\mathbf{A}^k.$$

Lema 1.5.3. Sean $p(x), q(x) \in \mathbb{C}[x]$ y $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces $p(\mathbf{A})$ y $q(\mathbf{A})$ conmutan, esto es, $p(\mathbf{A})q(\mathbf{A}) = q(\mathbf{A})p(\mathbf{A})$.

Demostración. Sean $p(x) = \sum\limits_{k=0}^{n} a_k x^k$ y $q(x) = \sum\limits_{j=0}^{m} b_j x^j$. Entonces

$$\begin{aligned} p(\mathbf{A})q(\mathbf{A}) &= \left(\sum_{k=0}^{n} a_k \mathbf{A}^k\right)\left(\sum_{j=0}^{m} b_j \mathbf{A}^j\right) = \sum_{k=0}^{n}\sum_{j=0}^{m}(a_k\mathbf{A}^k)(b_j\mathbf{A}^j) \\ &= \sum_{k=0}^{n}\sum_{j=0}^{m} a_k b_j \mathbf{A}^{k+j} = \sum_{j=0}^{m}\sum_{k=0}^{n} b_j a_k \mathbf{A}^{j+k} \\ &= \sum_{j=0}^{m}\sum_{k=0}^{n}(b_j\mathbf{A}^j)(a_k\mathbf{A}^k) = \left(\sum_{j=0}^{m} b_j \mathbf{A}^j\right)\left(\sum_{k=0}^{n} a_k \mathbf{A}^k\right) \\ &= q(\mathbf{A})p(\mathbf{A}). \end{aligned}$$

□

Teorema 1.5.4. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $p(x) \in \mathbb{C}[x]$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ y $\mathbf{x}$ un vector propio de $\mathbf{A}$ asociado a λ. Entonces $p(\lambda) \in \operatorname{Spec}(p(\mathbf{A}))$ y $\mathbf{x}$ es un vector propio de $p(\mathbf{A})$ asociado a $p(\lambda)$.

Demostración. Véase [9, página 45]. □

Teorema 1.5.5. Las matrices semejantes tienen el mismo polinomio característico. Por lo tanto, las matrices semejantes tienen los mismos valores propios con iguales multiplicidades algebraicas.

Demostración. Véase [9, página 45]. □

Finalizamos esta sección con el enunciado del Lema de Schur. La demostración del Lema de Schur puede ser consultada en [9, página 79]

Teorema 1.5.6 (Lema de Schur)**.** Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$ con valores propios $\lambda_1, \ldots, \lambda_n$ (los λ_i se repiten según sus multiplicidades algebraicas respectivas). Entonces $\mathbf{A}$ es unitariamente semejante a una matriz triangular superior cuyas entradas en su diagonal principal son los valores propios de $\mathbf{A}$ repetidos según sus multiplicidades algebraicas respectivas. Esto es, existen matrices $\mathbf{U}$ y $\mathbf{T}$, tales que $\mathbf{U}$ es unitaria, $\mathbf{T}$ es triangular superior y

$$\mathbf{A} = \mathbf{U}^H\mathbf{T}\mathbf{U}$$

En el lema de Schur, las entradas en la diagonal principal de la matriz triangular superior $\mathbf{T}$ son precisamente los valores propios de $\mathbf{A}$, incluyendo las repeticiones. El orden en que aparecen los valores propios en la diagonal de $\mathbf{T}$ depende de la matriz $\mathbf{U}$. Siempre se puede escoger la matriz $\mathbf{U}$ para que el orden en la diagonal principal de $\mathbf{T}$ sea cualquier orden prefijado.

Capítulo 2

Polinomios anuladores

Cuando se analiza una matriz $\mathbf{A} \in \mathbb{C}^{n \times n}$, una de las cosas más útiles de conocer es el conjunto de los polinomios que anulan a $\mathbf{A}$, es decir, el conjunto

$$\{p(x) \in \mathbb{C}[x] : p(\mathbf{A}) = 0\}.$$

Como se demostrará más adelante, este conjunto consta precisamente de los múltiplos de un único polinomio mónico fijo $q(x)$. Este polinomio $q(x)$ tiene la siguiente propiedad: si $p(x) \in \mathbb{C}[x]$, entonces $p(\mathbf{A}) = 0$ si y solo si $p(x) = s(x)q(x)$ para algún $s(x) \in \mathbb{C}[x]$. También se mostrará que el polinomio característico de $\mathbf{A}$, $P_{\mathbf{A}}(x)$, está en el conjunto de polinomios que anulan a $\mathbf{A}$, y más aún, que las raíces de $q(x)$ son precisamente los valores propios de $\mathbf{A}$.

2.1. Polinomios anuladores

Definición 2.1.1. Sea $\mathbf{A} \in \mathbb{C}^{n \times n}$. Un **polinomio anulador de A** es un polinomio $p(x) \in \mathbb{C}$ tal que $p(\mathbf{A}) = 0$. El conjunto formado por todos los polinomios anuladores de $\mathbf{A}$ será llamado **el anulador de A** y se denotará con el símbolo $\mathrm{Ann}(\mathbf{A})$. Así,

$$\mathrm{Ann}(\mathbf{A}) = \{p(x) \in \mathbb{C}[x] : p(\mathbf{A}) = 0\}.$$

A continuación demostramos el Teorema de Cayley - Hamilton, el cual garantiza que toda matriz $\mathbf{A}$ satisface su polinomio característico, y en particular, que $\mathrm{Ann}(\mathbf{A})$ es un conjunto no vacío.

Teorema 2.1.1 (Cayley - Hamilton)**.** Toda matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ satisface su polinomio característico, es decir,

$$P_{\mathbf{A}}(\mathbf{A}) = 0.$$

Demostración. Supongamos que

$$P_{\mathbf{A}}(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0,$$

donde $c_0, c_1, \ldots, c_{n-1} \in \mathbb{C}$. Por el lema de Schur (Teorema 1.5.6) existen $\mathbf{U}, \mathbf{T} \in \mathbb{C}^{n\times n}$ con $\mathbf{U}$ unitaria y $\mathbf{T}$ triangular superior tales que $\mathbf{A} = \mathbf{U}^H\mathbf{T}\mathbf{U}$. Usando el hecho de que $(\mathbf{U}^H\mathbf{T}\mathbf{U})^k = \mathbf{U}^H\mathbf{T}^k\mathbf{U}$ para todo $k \in \mathbb{Z}^+$, tenemos

$$\begin{aligned}
P_{\mathbf{A}}(\mathbf{A}) &= \mathbf{A}^n + c_{n-1}\mathbf{A}^{n-1} + \cdots + c_1\mathbf{A} + c_0\mathbf{I}_n \\
&= (\mathbf{U}^H\mathbf{T}\mathbf{U})^n + c_{n-1}(\mathbf{U}^H\mathbf{T}\mathbf{U})^{n-1} + \cdots + c_1(\mathbf{U}^H\mathbf{T}\mathbf{U}) + c_0(\mathbf{U}^H\mathbf{I}_n\mathbf{U}) \\
&= \mathbf{U}^H\mathbf{T}^n\mathbf{U} + \mathbf{U}^H c_{n-1}\mathbf{T}^{n-1}\mathbf{U} + \cdots + \mathbf{U}^H c_1\mathbf{T}\mathbf{U} + \mathbf{U}^H c_0\mathbf{I}_n\mathbf{U} \\
&= \mathbf{U}^H(\mathbf{T}^n + c_{n-1}\mathbf{T}^{n-1} + \cdots + c_1\mathbf{T} + c_0\mathbf{I}_n)\mathbf{U} \\
&= \mathbf{U}^H(P_{\mathbf{A}}(\mathbf{T}))\mathbf{U}.
\end{aligned}$$

Ahora, supongamos que $\lambda_1, \ldots, \lambda_n$ son los valores propios de $\mathbf{A}$ (en la lista $\lambda_1, \ldots, \lambda_n$ asumimos que cada valor propio de $\mathbf{A}$ aparece tantas veces como su multiplicidad algebraica) ordenados de tal manera que $\mathbf{T}_{ii} = \lambda_i$ para $i = 1, \ldots, n$. Luego,

$$\begin{aligned}
P_{\mathbf{A}}(x) &= (x - \lambda_1)(x - \lambda_2)\,\cdots\,(x - \lambda_n) \\
&= (x - \mathbf{T}_{11})(x - \mathbf{T}_{22})\,\cdots\,(x - \mathbf{T}_{nn}).
\end{aligned}$$

Esto implica que

$$P_{\mathbf{A}}(\mathbf{T}) = (\mathbf{T} - \mathbf{T}_{11}\mathbf{I}_n)(\mathbf{T} - \mathbf{T}_{22}\mathbf{I}_n)\,\cdots\,(\mathbf{T} - \mathbf{T}_{nn}\mathbf{I}_n).$$

Pero, la matriz $\mathbf{T} - \mathbf{T}_{11}\mathbf{I}_n$ tiene la primera columna nula, la matriz

$$(\mathbf{T} - \mathbf{T}_{11}\mathbf{I}_n)(\mathbf{T} - \mathbf{T}_{22}\mathbf{I}_n)$$

tiene las dos primeras columnas nulas, y así sucesivamente. De esta manera hallamos que la matriz

$$P_{\mathbf{A}}(\mathbf{T}) = (\mathbf{T} - \mathbf{T}_{11}\mathbf{I}_n)(\mathbf{T} - \mathbf{T}_{22}\mathbf{I}_n)\,\cdots\,(\mathbf{T} - \mathbf{T}_{nn}\mathbf{I}_n)$$

es la matriz nula, pues todas sus columnas son nulas. En consecuencia

$$P_{\mathbf{A}}(\mathbf{A}) = \mathbf{U}^H(P_{\mathbf{A}}(\mathbf{T}))\mathbf{U} = \mathbf{U}^H\, 0\, \mathbf{U} = 0.$$

□

Teorema 2.1.2. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces existe un único polinomio mónico no constante $q(x)$ de grado minimal en $\mathrm{Ann}(\mathbf{A})$.

Demostración. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Por el teorema de Cayley-Hamilton, $P_{\mathbf{A}}(x) \in \mathrm{Ann}(\mathbf{A})$, así que el conjunto

$$\mathcal{M} = \Big\{k \in \mathbb{Z}^+ : k = \deg(p(x)) \text{ para algún } p(x) \in \mathrm{Ann}(\mathbf{A}) - \{0\}\Big\}$$

es no vacío. Por el principio del buen orden, existe el mínimo de $\mathcal{M}$. Sea $m = \text{mín}\,\mathcal{M}$. Entonces $1 \leq m$ por la definición del conjunto $\mathcal{M}$, y también

$$m \leq \deg(P_{\mathbf{A}}(x)) = n,$$

puesto que $P_{\mathbf{A}}(x) \in \mathcal{M}$. Ahora, por la definición de $\mathcal{M}$, existe $p(x) \in \mathrm{Ann}(\mathbf{A})$ tal que $\deg(p(x)) = m$. Escribimos

$$p(x) = c_m x^m + \cdots + c_1 x + c_0$$

donde $c_m, \ldots, c_1, c_0 \in \mathbb{C}$ y $c_m \neq 0$. Si definimos $q(x) = (c_m)^{-1}p(x)$, entonces tenemos

$$q(\mathbf{A}) = (c_m)^{-1}p(\mathbf{A}) = (c_m)^{-1}0 = 0,$$

así que $q(x)$ es un polinomio mónico no constante de grado minimal en $\mathrm{Ann}(\mathbf{A})$.

Para mostrar la unicidad, supongamos que $t(x)$ es un polinomio mónico no constante de grado m en $\mathrm{Ann}(\mathbf{A})$. Si $t(x)$ no es múltiplo de $q(x)$, entonces por el algoritmo de la división existen $s(x)$, $r(x) \in \mathbb{C}\,[x]$, con $r(x) \neq 0$ y $\deg(r(x)) < m$, tales que

$$t(x) = s(x)q(x) + r(x).$$

Ahora,

$$r(\mathbf{A}) = t(\mathbf{A}) - s(\mathbf{A})q(\mathbf{A}) = 0 - s(\mathbf{A})0 = 0$$

así que $r(x) \in \text{Ann}(\mathbf{A})$. Esto es una contradicción, puesto que en $\text{Ann}(\mathbf{A})$ no hay polinomios no nulos de grado menor que m. Por lo tanto $t(x)$ es un múltiplo de $q(x)$. Luego, existe $u(x) \in \mathbb{C}[x]$ tal que $t(x) = u(x)q(x)$. Como

$$\deg(q(x)) = m = \deg(t(x)) = \deg(u(x)q(x)) = \deg(u(x)) + m,$$

se sigue que $u(x)$ es constante, y como $t(x)$ y $q(x)$ son mónicos, necesariamente $u(x) = 1$ y así $t(x) = q(x)$. □

2.2. Polinomio minimal

Definición 2.2.1. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. El único polinomio mónico de grado minimal en $\text{Ann}(\mathbf{A})$ será llamado el **polinomio minimal** de $\mathbf{A}$. Se usará el símbolo $m_{\mathbf{A}}(x)$ para representar al polinomio minimal de $\mathbf{A}$.

Teorema 2.2.1. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces

$$\text{Ann}(\mathbf{A}) = \{p(x)m_{\mathbf{A}}(x) : p(x) \in \mathbb{C}[x]\}. \tag{2.1}$$

Demostración. Sea $t(x) \in \text{Ann}(\mathbf{A})$ y supongamos que $t(x)$ no es múltiplo de $m_{\mathbf{A}}(x)$. Por el algoritmo de la división, existen $s(x), r(x)$ en $\mathbb{C}[x]$ tales que

$$t(x) = s(x)m_{\mathbf{A}}(x) + r(x),$$

con $r(x) \neq 0$ y $\deg(r(x)) < \deg(m_{\mathbf{A}}(x))$. Luego tenemos

$$r(\mathbf{A}) = t(\mathbf{A}) - s(\mathbf{A})m_{\mathbf{A}}(\mathbf{A}) = 0 - s(\mathbf{A})0 = 0$$

así que $r(x) \in \text{Ann}(\mathbf{A})$, lo cual es absurdo porque en $\text{Ann}(\mathbf{A})$ no existen polinomios no nulos de grado menor que el grado de $m_{\mathbf{A}}(x)$. Luego, $t(x)$ es múltiplo de $m_{\mathbf{A}}(x)$, y esto termina la prueba de la primera contenencia.
Para demostrar la contenencia contraria, sea $p(x) \in \mathbb{C}[x]$. Entonces

$$m_{\mathbf{A}}(\mathbf{A})\, p(\mathbf{A}) = 0\, p(\mathbf{A}) = 0,$$

por lo que $m_{\mathbf{A}}(x)\, p(x) \in \text{Ann}(\mathbf{A})$. □

Finalizamos este capítulo probando que para toda matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$,

$$\mathrm{Spec}(\mathbf{A}) = \{\alpha \in \mathbb{C} : \alpha \text{ es raíz de } m_{\mathbf{A}}(x)\},$$

y que la multiplicidad de cada valor propio λ de $\mathbf{A}$ como raíz de $m_{\mathbf{A}}(x)$ es menor o igual que la multiplicidad algebraica de λ.

Teorema 2.2.2. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda_1, \dots, \lambda_k$ los distintos valores propios de $\mathbf{A}$ con multiplicidades algebraicas $n_1, \dots, n_k$, respectivamente. Entonces

$$m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x - \lambda_i)^{m_i}$$

donde $1 \leq m_i \leq n_i$ para $i = 1, \dots, k$.

Demostración. Sea λ un valor propio de $\mathbf{A}$. Entonces existe $\mathbf{x} \in \mathbb{C}^n - \{0\}$ tal que $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Por el Teorema 1.5.4, $\mathbf{x}$ es un vector propio de $m_{\mathbf{A}}(\mathbf{A})$ asociado al valor propio $m_{\mathbf{A}}(\lambda)$. Como

$$m_{\mathbf{A}}(\lambda)\mathbf{x} = m_{\mathbf{A}}(\mathbf{A})\mathbf{x} = 0\mathbf{x} = 0$$

y $\mathbf{x} \neq 0$, resulta que $m_{\mathbf{A}}(\lambda) = 0$, es decir, λ es una raíz de $m_{\mathbf{A}}(x)$. Hemos mostrado así que toda raíz de $P_{\mathbf{A}}(x)$ (todo valor propio de $\mathbf{A}$) también es raíz de $m_{\mathbf{A}}(x)$. Por otro lado, como $P_{\mathbf{A}}(x) \in \mathrm{Ann}(\mathbf{A})$, entonces por el Teorema 2.2.1, $m_{\mathbf{A}}(x)$ divide a $P_{\mathbf{A}}(x)$. Luego, por el Teorema 1.4.6, toda raíz de $m_{\mathbf{A}}(x)$ es también raíz de $P_{\mathbf{A}}(x)$. Concluimos que $m_{\mathbf{A}}(x)$ y $P_{\mathbf{A}}(x)$ tienen exactamente las mismas raíces.

Finalmente, como $m_{\mathbf{A}}(x)$ divide a $P_{\mathbf{A}}(x)$, nuevamente por el Teorema 1.4.6, para cada raíz α de $m_{\mathbf{A}}(x)$, la multiplicidad de α como raíz de $m_{\mathbf{A}}(x)$ es menor o igual que la multiplicidad algebraica de α. Si m_i es la multiplicidad de λ_i como raíz de $m_{\mathbf{A}}(x)$, entonces tenemos que $1 \leq m_i \leq n_i$, $i = 1, \dots, k$, y por el Teorema 1.4.5, se tiene

$$m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x - \lambda_i)^{m_i} \tag{2.2}$$

$\square$

Capítulo 3

Vectores propios generalizados

En este capítulo se definen los conceptos de valor propio generalizado, subespacios propios generalizados, cadenas de Jordan y subespacios de Jordan asociados a una matriz dada $\mathbf{A} \in \mathbb{C}^{n\times n}$. Se demuestran resultados fuertes sobre descomposiciones de $\mathbb{C}^n$ como suma directa de subespacios propios generalizados asociados a $\mathbf{A}$, y de los subespacios propios generalizados como suma directa de espacios de Jordan asociados a la matriz $\mathbf{A}$. Estos resultados preparan el terreno para la demostración del teorema de Jordan, el cual demostraremos en el capítulo siguiente.

3.1. Vectores propios generalizados

Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$. Entonces $\mathcal{N}(\mathbf{A}-\lambda\mathbf{I}_n) \neq \{0\}$ y por el Lema 1.3.13, existe un entero positivo p tal que

$$\{\mathbf{0}\} \subsetneqq \cdots \subsetneqq \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p-1}) \subsetneqq \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p}) = \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p+k}),$$

para todo $k \in \mathbb{Z}^+$. Esto le da sentido a la siguiente definición.

Definición 3.1.1. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$. El **índice** de λ como valor propio de $\mathbf{A}$ es el entero positivo p tal que

$$\{\mathbf{0}\} \subsetneqq \cdots \subsetneqq \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p-1}) \subsetneqq \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p}) = \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^{p+k})$$

para todo $k \in \mathbb{Z}^+$. El índice de λ como valor propio de $\mathbf{A}$ lo denotaremos por el símbolo $i_{\mathbf{A}}(\lambda)$.

Una propiedad básica que podemos probar acerca de $i_{\mathbf{A}}(\lambda)$ se establece en el siguiente teorema.

Teorema 3.1.1. Sean $\mathbf{A} \in \mathbb{C}^{n\times n} - \{0\}$ y $\lambda \in \mathrm{Spec}(\mathbf{A})$. Entonces

$$\dim(\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{i_{\mathbf{A}}(\lambda)}) \geq i_{\mathbf{A}}(\lambda).$$

Demostración. Como $\{0\} \subsetneqq \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^1) \subsetneqq \cdots \subsetneqq \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{i_{\mathbf{A}}(\lambda)}$ y

$$1 \leq \dim(\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^1) < \cdots < \dim(\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{i_{\mathbf{A}}(\lambda)}),$$

entonces $\dim(\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{i_{\mathbf{A}}(\lambda)}) \geq i_{\mathbf{A}}(\lambda)$. □

Definición 3.1.2. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \mathrm{Spec}(\mathbf{A})$ tal que $i_{\mathbf{A}}(\lambda) = p$. Para cada $r \in \{1, \ldots, p\}$, el subespacio $\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^r)$ es llamado el **subespacio propio generalizado de A de orden r asociado al valor propio** λ. Si

$$\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^r) - \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-1}),$$

entonces $\mathbf{x}$ es llamado **vector propio generalizado de A de orden r asociado al valor propio** λ.

Si $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \mathrm{Spec}(\mathbf{A})$, entonces el subespacio propio generalizado de $\mathbf{A}$ de orden 1 asociado al valor propio λ no es más que el subespacio propio de $\mathbf{A}$ asociado al valor propio λ, y los vectores propios generalizados de $\mathbf{A}$ de orden 1 asociados al valor propio λ son precisamente los vectores propios de $\mathbf{A}$ asociados al valor propio λ, pues $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = 0$ es equivalente a $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

Definición 3.1.3. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathrm{Spec}(\mathbf{A})$ con $i_{\mathbf{A}}(\lambda) = p$, y $\mathbf{x}_r$ un vector propio generalizado de $\mathbf{A}$ de orden r, donde $1 \leq r \leq p$. Una lista de vectores

$$\mathbf{x}_r, \mathbf{x}_{r-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1,$$

donde

$$\mathbf{x}_{r-k} = (\mathbf{A} - \lambda\mathbf{I}_n)^k\mathbf{x}_r, \quad k = 0, 1, \ldots, r-1,$$

es llamada una **cadena de Jordan de A** de longitud r asociada a λ. El subespacio

$$\mathrm{gen}\,\{\mathbf{x}_1, \ldots, \mathbf{x}_{r-1}, \mathbf{x}_r\}$$

es llamado un **subespacio cíclico de Jordan** o simplemente **subespacio de Jordan**.

Con la notación e información contenida en la Definición 3.1.3, tenemos las siguientes observaciones:

1. En una cadena de Jordan $\mathbf{x}_r, \mathbf{x}_{r-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1$, se tiene $x_1 \neq 0$, ya que en caso contrario tendríamos
$$(\mathbf{A} - \lambda \mathbf{I}_n)^{r-1} \mathbf{x}_r = \mathbf{x}_1 = 0,$$
esto es, $\mathbf{x}_r \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1})$, lo cual es absurdo, pues $\mathbf{x}_r$ es un vector propio generalizado de $\mathbf{A}$ de orden r asociado a λ.

2. Tenemos
$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_1 = (\mathbf{A} - \lambda \mathbf{I}_n)(\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}\mathbf{x}_r = (\mathbf{A} - \lambda \mathbf{I}_n)^r \mathbf{x}_r = 0.$$
Luego, $\mathbf{x}_1 \in \mathcal{N}(\mathbf{A} - \lambda \mathbf{I}_n) - \{0\}$ y esto quiere decir que $\mathbf{x}_1$ es un vector propio generalizado de $\mathbf{A}$ de orden 1 asociado a λ, lo cual equivale a que $\mathbf{x}_1$ es un vector propio de $\mathbf{A}$ asociado a λ.

3. Por otro lado
$$(\mathbf{A} - \lambda \mathbf{I}_n)^2\mathbf{x}_2 = (\mathbf{A} - \lambda \mathbf{I}_n)^2(\mathbf{A} - \lambda \mathbf{I}_n)^{r-2}\mathbf{x}_r = (\mathbf{A} - \lambda \mathbf{I}_n)^r \mathbf{x}_r = 0$$
y
$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_2 = (\mathbf{A} - \lambda \mathbf{I}_n)(\mathbf{A} - \lambda \mathbf{I}_n)^{r-2}\mathbf{x}_r = (\mathbf{A} - \lambda \mathbf{I}_n)^{r-1} \mathbf{x}_r \neq 0.$$
Esto muestra que $\mathbf{x}_2 \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^2) - \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n))$, lo que significa que $\mathbf{x}_2$ es un vector propio generalizado de $\mathbf{A}$ de orden 2 asociado a λ.

4. Continuando este proceso inductivamente, vemos que para todo $i \in \{1, \ldots, r\}$, $\mathbf{x}_i$ es un vector propio generalizado de $\mathbf{A}$ de orden i asociado a λ.

Resumimos estas observaciones en el siguiente teorema.

Teorema 3.1.2. Sean $\mathbf{A} \in \mathbb{C}^{n \times n}$ y $\mathbf{x}_r, \mathbf{x}_{r-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1$, una cadena de Jordan de $\mathbf{A}$ de longitud r asociada al valor propio λ de $\mathbf{A}$. Entonces para todo $i \in \{1, \ldots, r\}$, $\mathbf{x}_i$ es un vector propio generalizado de $\mathbf{A}$ de orden i asociado a λ.

Observemos también que para una cadena de Jordan $\mathbf{x}_r, \mathbf{x}_{r-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1$, de $\mathbf{A}$ asociada a λ, tenemos que para todo $i \in \{2, \ldots, r\}$,

$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_i = (\mathbf{A} - \lambda \mathbf{I}_n)(\mathbf{A} - \lambda \mathbf{I}_n)^{r-i}\mathbf{x}_r = (\mathbf{A} - \lambda \mathbf{I}_n)^{r-i+1}\mathbf{x}_r = \mathbf{x}_{i-1},$$

o equivalentemente

$$\mathbf{A}\mathbf{x}_i - \lambda \mathbf{x}_i = \mathbf{x}_{i-1}$$

Conviniendo que $\mathbf{x}_0 = 0$, podemos escribir

$$\mathbf{A}\mathbf{x}_i = \lambda \mathbf{x}_i + \mathbf{x}_{i-1},$$

para todo $i \in \{1, 2, \ldots, r\}$. Concluimos que el conjunto de ecuaciones

$$(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_i = \mathbf{x}_{i-1}, \quad i = 1, 2, \ldots, r$$

es equivalente al conjunto de ecuaciones

$$\mathbf{A}\mathbf{x}_i = \lambda \mathbf{x}_i + \mathbf{x}_{i-1}, \quad i = 1, 2, \ldots, r.$$

Nótese que $\mathbf{x}_1$ es solución no nula del sistema de ecuaciones lineales $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} = 0$ y para i variando de 2 hasta r tenemos que $\mathbf{x}_i$ es solución del sistema de ecuaciones lineales $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_i = \mathbf{x}_{i-1}$.

En el resto de esta sección demostraremos varias propiedades sobre las cadenas de Jordan.

Lema 3.1.3. Sean $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ con $i_{\mathbf{A}}(\lambda) = p$ y $\mathbf{x}_r$ un vector propio generalizado de $\mathbf{A}$ de orden r, $1 \leq r \leq p$. Entonces la cadena de Jordan

$$\mathbf{x}_r, \mathbf{x}_{r-1}, \ldots, \mathbf{x}_2, \mathbf{x}_1,$$

donde

$$\mathbf{x}_{r-k} = (\mathbf{A} - \lambda \mathbf{I}_n)^k \mathbf{x}_r, \quad k = 0, 1, \ldots, r-1,$$

es un conjunto de vectores linealmente independiente.

Demostración. Sean $c_r, \ldots, c_1 \in \mathbb{C}$ tales que

$$\sum_{i=1}^{r} c_i \mathbf{x}_i = 0.$$

Multiplicando ambos lados de esta ecuación por $(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}$ obtenemos

$$(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\sum_{i=1}^{r} c_i\mathbf{x}_i = (\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}0 = 0.$$

Ahora, teniendo en cuenta que para $i = 1, 2, \dots, r-1$,

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_i &= (\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}(\mathbf{A}-\lambda\mathbf{I}_n)^{r-i}\mathbf{x}_r\\ &= (\mathbf{A}-\lambda\mathbf{I}_n)^{r-i-1}(\mathbf{A}-\lambda\mathbf{I}_n)^{r}\mathbf{x}_r\\ &= (\mathbf{A}-\lambda\mathbf{I}_n)^{r-i-1}0\\ &= 0\end{aligned}$$

obtenemos

$$\begin{aligned}(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\sum_{i=1}^{r} c_i\mathbf{x}_i &= \sum_{i=1}^{r} c_i(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_i\\ &= c_r(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_r + \sum_{i=1}^{r-1} c_i(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_i\\ &= c_r(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_r + \sum_{i=1}^{r-1} c_i 0\\ &= c_r(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_r,\end{aligned}$$

de donde concluimos que $c_r = 0$, puesto que $(\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}\mathbf{x}_r \neq 0$. Así quedamos con la relación,

$$\sum_{i=1}^{r-1} c_i\mathbf{x}_i = 0.$$

Por un argumento similar al anterior, pero esta vez multiplicando por $(\mathbf{A}-\lambda\mathbf{I}_n)^{r-2}$ a ambos lados, hallamos que $c_{r-1} = 0$, y quedamos con la relación

$$\sum_{i=1}^{r-2} c_i\mathbf{x}_i = 0.$$

Continuando con este proceso llegamos a que $c_r = c_{r-1} = \cdots = c_1 = 0$. Se concluye así que el conjunto $\{\mathbf{x}_r, \dots, \mathbf{x}_1\}$ es linealmente independiente. □

Lema 3.1.4. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ y $\mathbf{x}_r, \mathbf{x}_{r-1}, \dots, \mathbf{x}_1$ una cadena de Jordan asociada a λ. Entonces todo vector propio de $\mathbf{A}$ que pertenezca al subespacio de Jordan gen $\{\mathbf{x}_r, \dots, \mathbf{x}_1\}$, es un múltiplo escalar de $\mathbf{x}_1$.

Demostración. Sea $\mathbf{x}$ un vector propio de $\mathbf{A}$. Entonces existe un valor propio α de $\mathbf{A}$ tal que $\mathbf{x}$ está asociado a α. Supongamos que $\mathbf{x} \in \operatorname{gen}\{\mathbf{x}_r, \ldots, \mathbf{x}_1\}$ y escribamos $\mathbf{x} = \sum_{i=1}^{r} c_i \mathbf{x}_i$, donde $c_i \in \mathbb{C}$, $i = 1, 2, \ldots, r$. Entonces

$$\mathbf{A}\mathbf{x} = \alpha\mathbf{x} = \alpha \sum_{i=1}^{r} c_i \mathbf{x}_i = \sum_{i=1}^{r} \alpha c_i \mathbf{x}_i.$$

Por otro lado, como $\mathbf{x}_r, \ldots, \mathbf{x}_1$ es una cadena de Jordan asociada a λ, tenemos

$$\begin{aligned}
\mathbf{A}\mathbf{x}_1 &= \lambda \mathbf{x}_1 \\
\mathbf{A}\mathbf{x}_2 &= \lambda \mathbf{x}_2 + \mathbf{x}_1 \\
&\vdots \\
\mathbf{A}\mathbf{x}_r &= \lambda \mathbf{x}_r + \mathbf{x}_{r-1}
\end{aligned}$$

Haciendo $\mathbf{x}_0 = 0$ y $c_{r+1} = 0$ podemos escribir

$$\begin{aligned}
\sum_{i=1}^{r} \alpha c_i \mathbf{x}_i = \mathbf{A}\mathbf{x} = \mathbf{A} \sum_{i=1}^{r} c_i \mathbf{x}_i = \sum_{i=1}^{r} c_i \mathbf{A}\mathbf{x}_i &= \sum_{i=1}^{r} c_i (\lambda \mathbf{x}_i + \mathbf{x}_{i-1}) \\
&= \sum_{i=1}^{r} c_i \lambda \mathbf{x}_i + \sum_{i=1}^{r} c_i \mathbf{x}_{i-1} = \sum_{i=1}^{r} c_i \lambda \mathbf{x}_i + \sum_{i=0}^{r-1} c_{i+1} \mathbf{x}_i \\
&= \sum_{i=1}^{r} c_i \lambda \mathbf{x}_i + \sum_{i=1}^{r} c_{i+1} \mathbf{x}_i.
\end{aligned}$$

Luego,

$$\sum_{i=1}^{r} (\lambda c_i + c_{i+1} - \alpha c_i) \mathbf{x}_i = 0.$$

Como $\{\mathbf{x}_r, \ldots, \mathbf{x}_1\}$ es linealmente independiente, se sigue que

$$\lambda c_i + c_{i+1} - \alpha c_i = 0,$$

para todo $i \in \{1, \ldots, r\}$. Dado que $c_{r+1} = 0$, tenemos que

$$\lambda c_r = \alpha c_r \tag{3.1}$$

y

$$\lambda c_i + c_{i+1} = \alpha c_i \tag{3.2}$$

para todo $i \in \{1, \ldots, r-1\}$. Afirmamos que $\alpha = \lambda$. En efecto, razonando por reducción al absurdo, supongamos que $\alpha \neq \lambda$. Luego por (3.1) se sigue que $c_r = 0$. Ahora, por (3.2), $\lambda c_{r-1} + 0 = \alpha c_{r-1}$, de donde $c_{r-1} = 0$. Nuevamente, por (3.2), $\lambda c_{r-2} + 0 = \alpha c_{r-2}$, y así $c_{r-2} = 0$. Siguiendo este proceso inductivo llegamos a que

$$c_r = c_{r-1} = \cdots = c_1 = 0,$$

lo cual implica que $\mathbf{x} = 0$, y esto es absurdo pues $\mathbf{x}$ es un vector propio de $\mathbf{A}$ asociado al valor propio α. Esto demuestra que $\alpha = \lambda$.

Finalmente, como $\alpha = \lambda$ y $\lambda c_i + c_{i+1} = \alpha c_i$, para todo $i \in \{1, \ldots, r-1\}$, obtenemos que $c_{i+1} = 0$ para todo $i \in \{1, \ldots, r-1\}$. Por lo tanto,

$$\mathbf{x} = \sum_{i=1}^{r} c_i \mathbf{x}_i = c_1 \mathbf{x}_1,$$

lo cual termina la demostración. $\square$

El Lema 3.1.4 dice que el subespacio de Jordan gen $\{\mathbf{x}_r, \ldots, \mathbf{x}_1\}$ asociado al valor propio λ, solo intersecta al subespacio propio $E_{\mathbf{A}}(\lambda)$ y además

$$E_{\mathbf{A}}(\lambda) \cap \operatorname{gen}\{\mathbf{x}_r, \ldots, \mathbf{x}_1\} = \{\alpha \mathbf{x}_1 : \alpha \in \mathbb{C}\}.$$

Lema 3.1.5. Sean $\mathbf{A} \in \mathbb{C}^{n \times n} - \{0\}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$ de índice p. Entonces para todo $r \in \{1, \ldots, p\}$ tenemos que

$$\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^r) \text{ si y solo si } (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}).$$

Demostración. Note que $\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^r)$ si y solo si $(\mathbf{A} - \lambda \mathbf{I}_n)^r \mathbf{x} = \mathbf{0}$. Esto es, $(\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}[(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}] = \mathbf{0}$, lo que equivale a que $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1})$. Por lo tanto, $\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^r)$ si, y solo si $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1})$. $\square$

Lema 3.1.6. Sean $\mathbf{A} \in \mathbb{C}^{n \times n}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$ de índice p. Entonces para todo $r \in \{2, \ldots, p\}$ tenemos que $\mathbf{x}$ es un vector propio generalizado de $\mathbf{A}$ de orden r asociado a λ si y solo si $(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}$ es un vector propio generalizado de $\mathbf{A}$ de orden $r-1$ asociado a λ.

Demostración. Por definición, que $\mathbf{x}$ sea un vector propio generalizado de $\mathbf{A}$ de orden r asociado a λ significa que $\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^r) - \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-1})$, o sea que $\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^r)$ y $\mathbf{x} \notin \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-1})$. Esto es equivalente a tener $(\mathbf{A} - \lambda\mathbf{I}_n)^r\mathbf{x} = \mathbf{0}$ y $(\mathbf{A} - \lambda\mathbf{I}_n)^{r-1}\mathbf{x} \neq \mathbf{0}$, que a su vez podemos reescribir de la siguiente forma: $(\mathbf{A} - \lambda\mathbf{I}_n)^{r-1}(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ y $(\mathbf{A} - \lambda\mathbf{I}_n)^{r-2}(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} \neq \mathbf{0}$. Esto último quiere decir que $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-1})$ y $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} \notin \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-2})$, o en otras palabras, que $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-1}) - \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^{r-2})$, y esto significa que $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x}$ es un vector propio generalizado de $\mathbf{A}$ de orden $r - 1$ asociado a λ. $\square$

3.2. Una descomposición de $\mathbb{C}^n$ como suma directa de subespacios propios generalizados

Los resultados de esta sección son pasos importantes para la demostración del teorema de Jordan. En el siguiente teorema mostramos que dada una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbb{C}^n$ puede descomponerse como una suma directa de subespacios propios generalizados de $\mathbf{A}$. Más adelante mostramos que todo subespacio propio generalizado admite una descomposición como suma directa de subespacios de Jordan asociados a $\mathbf{A}$.

Teorema 3.2.1. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\operatorname{Spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_k\}$ y $m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x - \lambda_i)^{m_i}$. Entonces

$$\mathbb{C}^n = \bigoplus_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i\mathbf{I}_n)^{m_i}).$$

Demostración. Primero mostremos que $\mathbb{C}^n = \sum_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i\mathbf{I}_n)^{m_i})$. En efecto, consideremos los k polinomios

$$q_j = \prod_{\substack{i=1 \\ i\neq j}}^{k}(x - \lambda_i)^{m_i}, \quad j = 1, 2, \ldots, k.$$

Notemos que $(x - \lambda_j)^{m_j}q_j(x) = m_{\mathbf{A}}(x)$, $j = 1, 2, \ldots, k$.

Estos polinomios $q_j(x)$ no tienen factores comunes excepto constantes. En efecto, si los polinomios $q_j(x)$, $j = 1, 2, \ldots, k$ tuvieran un factor común no constante, entonces tendrían un factor común lineal de la forma $x - \alpha$. Así, puesto que $x - \alpha$ divide a $q_1(x)$, tendríamos que $x-\alpha = x-\lambda_{j_1}$ para algún $j_1 \neq 1$, y a la vez, como $(x-\alpha) \mid q_{j_1}(x)$, resultaría que $x-\alpha = x-\lambda_{j_2}$ para algún $j_2 \neq j_1$, pero tendríamos así que $x-\lambda_{j_1} = x-\lambda_{j_2}$, de donde $\lambda_{j_1}=\lambda_{j_2}$, lo cual es absurdo.

Luego, por el Teorema 1.4.2, existen polinomios $s_j(x) \in \mathbb{C}[x]$, $j=1,2,\ldots,k$, tales que

$$1 = \sum_{j=1}^{k} q_j(x)s_j(x).$$

Se sigue que

$$\mathbf{I}_n = \sum_{j=1}^{k} q_j(\mathbf{A})s_j(\mathbf{A}).$$

Definamos $\mathbf{x}_j = (q_j(\mathbf{A})s_j(\mathbf{A}))\mathbf{x}$, $j=1,2,\ldots,k$. Entonces tenemos que

$$\mathbf{x} = \mathbf{I}_n\mathbf{x} = \left(\sum_{j=1}^{k} q_j(\mathbf{A})s_j(\mathbf{A})\right)\mathbf{x} = \sum_{j=1}^{k} q_j(\mathbf{A})s_j(\mathbf{A})\mathbf{x} = \sum_{j=1}^{k}\mathbf{x}_j.$$

Afirmamos que $\mathbf{x}_j \in \mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j})$ para todo $j \in \{1,2,\ldots,k\}$. En efecto, esto se sigue de lo siguiente:

$$\begin{aligned}
(\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}\mathbf{x}_j &= (\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}(q_j(\mathbf{A})s_j(\mathbf{A}))\mathbf{x}\\
&= m_{\mathbf{A}}(\mathbf{A})s_j(\mathbf{A})\mathbf{x}\\
&= 0s_j(\mathbf{A})\mathbf{x}\\
&= 0
\end{aligned}$$

Hemos mostrado que todo $\mathbf{x} \in \mathbb{C}^n$ puede escribirse como una suma

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k,$$

donde $\mathbf{x}_j \in \mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j})$ para todo $j \in \{1,2,\ldots,k\}$, esto es,

$$\mathbb{C}^n = \sum_{i=1}^{k} \mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}).$$

Ahora mostremos que esta suma es directa. En efecto, sean $j \in \{1,\ldots,k\}$ y

$$\mathbf{x} \in \mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}) \bigcap \sum_{\substack{i=1\\ i\neq j}}^{k} \mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}).$$

Entonces $(\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}\mathbf{x}=0$ y $\mathbf{x}=\sum\limits_{r=1,r\neq j}^{k}\mathbf{x}_r$, donde $\mathbf{x}_r\in\mathcal{N}((\mathbf{A}-\lambda_r\mathbf{I}_n)^{m_r})$ para todo $r\in\{1,\ldots,k\}-\{j\}$. Ahora, como $(x-\lambda_j)^{m_j}$ y $\prod\limits_{i=1,i\neq j}^{k}(x-\lambda_i)^{m_i}$ son primos relativos, por el Teorema 1.4.2, existen $s(x),t(x)\in\mathbb{C}[x]$ tales que

$$1=s(x)(x-\lambda_j)^{m_j}+t(x)\prod_{\substack{i=1\\ i\neq j}}^{k}(x-\lambda_i)^{m_i}.$$

Luego,

$$\begin{aligned}
\mathbf{x}&=\mathbf{I}_n\mathbf{x}\\
&=\left(s(\mathbf{A})(\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}+t(\mathbf{A})\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\right)\mathbf{x}\\
&=s(\mathbf{A})(\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}\mathbf{x}+t(\mathbf{A})\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\mathbf{x}\\
&=s(\mathbf{A})(0)+t(\mathbf{A})\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\left(\sum_{\substack{r=1\\ r\neq j}}^{k}\mathbf{x}_r\right)\\
&=t(\mathbf{A})\sum_{\substack{r=1\\ r\neq j}}^{k}\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\mathbf{x}_r.
\end{aligned}$$

Por otro lado tenemos que, para todo $r\in\{1,\ldots,k\}-\{j\}$,

$$\begin{aligned}
\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\mathbf{x}_r&=\left(\prod_{\substack{i=1\\ i\neq j,r}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\right)(\mathbf{A}-\lambda_r\mathbf{I}_n)^{m_r}\mathbf{x}_r\\
&=\prod_{\substack{i=1\\ i\neq j,r}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}(0)\\
&=0.
\end{aligned}$$

Luego,

$$\mathbf{x}=t(\mathbf{A})\sum_{\substack{r=1\\ r\neq j}}^{k}\prod_{\substack{i=1\\ i\neq j}}^{k}(\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}\mathbf{x}_r=t(\mathbf{A})\sum_{\substack{r=1\\ r\neq j}}^{k}0=t(\mathbf{A})(0)=0.$$

Esto demuestra que

$$\mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j})\bigcap\sum_{\substack{i=1\\ i\neq j}}^{k}\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i})=\{0\}.$$

Por el Teorema 1.3.6 concluimos que

$$\mathbb{C}^n = \bigoplus_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{m_i}).$$

$\square$

Antes de proceder a demostrar que todo subespacio propio generalizado se descompone como suma directa de subespacios de Jordan, probamos el siguiente lema.

Lema 3.2.2. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ y $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_s\}$ un conjunto linealmente independiente formado por vectores propios generalizados de $\mathbf{A}$ de orden r asociados a λ, donde $r \geq 2$, de tal manera que

$$\mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^r) = \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}) \oplus \operatorname{gen}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_s\}.$$

Entonces el conjunto

$$\{(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_1, (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_2, \ldots, (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_s\},$$

formado por vectores propios generalizados de $\mathbf{A}$ de orden $r-1$ asociados a λ, también es linealmente independiente.

Demostración. Sean $c_1, c_2, \ldots, c_s \in \mathbb{C}$ tales que

$$c_1(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_1 + c_2(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_2 + \cdots + c_s(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_s = 0.$$

De aquí se sigue que

$$(\mathbf{A} - \lambda \mathbf{I}_n)(c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_s\mathbf{y}_s) = 0,$$

así que multiplicando ambos lados de esta última ecuación por $(\mathbf{A}-\lambda \mathbf{I}_n)^{r-2}$ obtenemos

$$(\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}(c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_s\mathbf{y}_s) = (\mathbf{A} - \lambda \mathbf{I}_n)^{r-2}0 = 0.$$

Luego, $c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_s\mathbf{y}_s \in \mathcal{N}((\mathbf{A} - \lambda \mathbf{I}_n)^{r-1}) \cap \operatorname{gen}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_s\}$. La hipótesis del lema implica que

$$c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \cdots + c_s\mathbf{y}_s = 0.$$

Como $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_s\}$ es linealmente independiente tenemos que $c_1 = c_2 = \cdots = c_s = 0$. Esto demuestra que el conjunto $\{(\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_1, (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_2, \ldots, (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{y}_s\}$ es linealmente independiente. $\square$

Terminamos la sección demostrando que para toda matiz $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$ de índice p, el subespacio propio generalizado $\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p)$ se puede expresar como una suma directa de subespacios de Jordan asociados a λ.

Teorema 3.2.3. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\lambda \in \operatorname{Spec}(\mathbf{A})$ con $i_{\mathbf{A}}(\lambda) = p$. Entonces existe una descomposición

$$\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p) = (\mathcal{G}_p^{(1)} \oplus \cdots \oplus \mathcal{G}_p^{(t_p)}) \oplus \cdots \oplus (\mathcal{G}_1^{(t_p+\cdots+t_2+1)} \oplus \cdots \oplus \mathcal{G}_1^{(t_p+\cdots+t_2+t_1)})$$

donde cada $\mathcal{G}_j^{(i)}$ es un subespacio cíclico de Jordan de dimensión j.

Demostración. El siguiente diagrama servirá para guiarse a lo largo de la prueba y comprender su construcción (donde $t = t_p + t_{p-1} + \cdots + t_1$):

$$\begin{array}{ccccc}
\mathbf{x}_1^{(1)}, & \mathbf{x}_2^{(1)}, & \ldots, & \mathbf{x}_{p-1}^{(1)}, & \mathbf{x}_p^{(1)}, \\
\vdots & \vdots & & \vdots & \vdots \\
\mathbf{x}_1^{(t_p)}, & \mathbf{x}_2^{(t_p)}, & \ldots, & \mathbf{x}_{p-1}^{(t_p)}, & \mathbf{x}_p^{(t_p)}, \\
\mathbf{x}_1^{(t_p+1)}, & \mathbf{x}_2^{(t_p+1)}, & \ldots, & \mathbf{x}_{p-1}^{(t_p+1)}, & \\
\vdots & \vdots & & \vdots & \\
\mathbf{x}_1^{(t_p+t_{p-1})}, & \mathbf{x}_2^{(t_p+t_{p-1})}, & \cdots, & \mathbf{x}_{p-1}^{t_p+t_{p-1}}, & \\
\vdots & \vdots & & & \\
\mathbf{x}_1^{(t-t_1)}, & \mathbf{x}_2^{(t-t_1)}, & & & \\
\mathbf{x}_1^{(t-t_1+1)}, & & & & \\
\vdots & & & & \\
\mathbf{x}_1^{(t)}, & & & &
\end{array} \tag{3.3}$$

Definimos $\mathcal{U}_i = \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^i)$, para $i = 1, 2, \ldots, p$. Como $\mathcal{U}_{p-1} \subsetneqq \mathcal{U}_p$, podemos hallar vectores linealmente independientes

$$\mathbf{x}_p^{(1)}, \ldots, \mathbf{x}_p^{(t_p)}$$

en $\mathcal{U}_p - \mathcal{U}_{p-1}$ tales que

$$\mathcal{U}_p = \mathcal{U}_{p-1} \oplus \operatorname{gen}\left\{\mathbf{x}_p^{(1)}, \ldots, \mathbf{x}_p^{(t_p)}\right\}.$$

Estos vectores $\mathbf{x}_p^{(1)}, \ldots, \mathbf{x}_p^{(t_p)}$ son los de la última columna a la derecha en (3.3), y son vectores propios generalizados de $\mathbf{A}$ de orden p asociados al valor propio λ. Por los Lemas 3.1.6 y 3.2.2, los vectores

$$\mathbf{x}_{p-1}^{(1)} = (\mathbf{A} - \lambda \mathbf{I}_n)(\mathbf{x}_p^{(1)}), \ldots, \mathbf{x}_{p-1}^{(t_p)} = (\mathbf{A} - \lambda \mathbf{I}_n)(\mathbf{x}_p^{(t_p)})$$

forman un conjunto linealmente independiente formado por vectores propios generalizados de $\mathbf{A}$ de orden $p-1$ asociados a λ.

Observemos que $\mathcal{U}_{p-2} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-1}^{(1)}, \ldots, \mathbf{x}_{p-1}^{(t_p)}\right\} \subseteq \mathcal{U}_{p-1}$. Podemos escoger vectores linealmente independientes $\mathbf{x}_{p-1}^{(t_p+1)}, \ldots, \mathbf{x}_{p-1}^{(t_p+t_{p-1})}$ en $\mathcal{U}_{p-1} - \mathcal{U}_{p-2}$ (siempre que $\mathcal{U}_{p-2} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-1}^{(1)}, \ldots, \mathbf{x}_{p-1}^{(t_p)}\right\} \subsetneq \mathcal{U}_{p-1}$, pues en caso de igualdad tomamos $t_{p-1} = 0$) tales que

$$\begin{aligned}\mathcal{U}_{p-1} &= \mathcal{U}_{p-2} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-1}^{(1)}, \ldots, \mathbf{x}_{p-1}^{(t_p)}\right\} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-1}^{(t_p+1)}, \ldots, \mathbf{x}_{p-1}^{(t_p+t_{p-1})}\right\} \\ &= \mathcal{U}_{p-2} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-1}^{(1)}, \ldots, \mathbf{x}_{p-1}^{(t_p)}, \mathbf{x}_{p-1}^{(t_p+1)}, \ldots, \mathbf{x}_{p-1}^{(t_p+t_{p-1})}\right\}.\end{aligned}$$

El conjunto de vectores

$$\mathbf{x}_{p-1}^{(1)}, \ldots, \mathbf{x}_{p-1}^{(t_p)}, \mathbf{x}_{p-1}^{(t_p+1)}, \ldots, \mathbf{x}_{p-1}^{(t_p+t_{p-1})}$$

es linealmente independiente (y forman la penúltima columna a la derecha en (3.3)).

Supongamos que ya se ha construido la columna $k-$ésima de derecha a izquierda en el diagrama (3.3), donde $k \in \{1, \ldots, p-1\}$, donde tal columna $k-$ésima está formada por vectores propios generalizados de orden $p-k+1$ asociados a λ (es decir, están en $\mathcal{U}_{p-k+1} - \mathcal{U}_{p-k}$):

$$\mathbf{x}_{p-k+1}^{(1)}, \ldots, \mathbf{x}_{p-k+1}^{(t_p)}, \ldots, \mathbf{x}_{p-k+1}^{(t_p+\cdots+t_{p-k+2}+1)}, \ldots, \mathbf{x}_{p-k+1}^{(t_p+\cdots+t_{p-k+1})}$$

que forman un conjunto linealmente independiente en $\mathcal{U}_{p-k+1} - \mathcal{U}_{p-k}$. Definimos

$$\mathbf{x}_{p-k}^{(j)} = (\mathbf{A} - \lambda \mathbf{I}_n)\mathbf{x}_{p-k+1}^{(j)}$$

para $j = 1, 2, \ldots, t_p + \cdots + t_{p-k+1}$. Por los Lemas 3.1.6 y 3.2.2, los vectores $\mathbf{x}_{p-k}^{(j)}$, $j = 1, 2, \ldots, t_p + \cdots + t_{p-k+1}$ son vectores propios generalizados de orden $p-k$ asociados a λ. Además se tiene que

$$\mathcal{U}_{p-k-1} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-k}^{(1)}, \ldots, \mathbf{x}_{p-k}^{(t_p)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+2}+1)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1})}\right\} \subseteq \mathcal{U}_{p-k}$$

Si se da la igualdad en la contenencia anterior escogemos $t_{p-k} = 0$, pero si la contenencia es estricta, se pueden escoger vectores

$$\mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+1)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+t_{p-k})}$$

de tal forma que $\mathcal{U}_{p-k}$ es igual a

$$\mathcal{U}_{p-k-1)} \oplus \operatorname{gen}\left\{\mathbf{x}_{p-k}^{(1)}, \ldots, \mathbf{x}_{p-k}^{(t_p)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+1)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+t_{p-k})}\right\}$$

Los vectores

$$\mathbf{x}_{p-k}^{(1)}, \ldots, \mathbf{x}_{p-k}^{(t_p)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+1)}, \ldots, \mathbf{x}_{p-k}^{(t_p+\cdots+t_{p-k+1}+t_{p-k})}$$

forman la columna $(k+1)-$ésima de izquierda a derecha en (3.3), y por construcción, son vectores propios generalizados de orden $p-k$ asociados a λ.

Así que tenemos p columnas en el diagrama (3.3), la columna $j-$ésima de izquierda a derecha está formada por vectores propios generalizados de orden j. Además, cada fila del diagrama (3.3) es una cadena de Jordan, cuya longitud está entre 1 y p. Ahora, para cada $j \in \{1, \ldots, p\}$, sea

$$\mathcal{C}_j = \left\{\mathbf{x}_j^{(1)}, \ldots, \mathbf{x}_j^{(t_p+\cdots+t_j))}\right\}.$$

Es decir, $\mathcal{C}_j$ es el conjunto formado por los vectores en la columna $j-$ésima del diagrama (3.3).

Si $j_1 \neq j_2$, entonces $\mathcal{C}_{j_1} \cap \mathcal{C}_{j_2} = \varnothing$ puesto que $\mathcal{C}_{j_1}$ está formado por vectores propios generalizados de orden j_1, mientras que $\mathcal{C}_{j_2}$ está formado por vectores propios generalizados de orden j_2. Cada $\mathcal{C}_j$ es un conjunto linealmente independiente por construcción. Ahora, $\mathcal{C}_1$ es una base para $\mathcal{U}_1$, así que $\mathcal{U}_1 = \operatorname{gen}(\mathcal{C}_1)$, y tenemos

$$\mathcal{U}_2 = \operatorname{gen}(\mathcal{C}_1) \oplus \operatorname{gen}(\mathcal{C}_2),$$

$$\vdots$$

$$\mathcal{U}_{p-1} = \operatorname{gen}(\mathcal{C}_1) \oplus \cdots \oplus \operatorname{gen}(\mathcal{C}_{p-1})$$

y

$$\mathcal{U}_p = \operatorname{gen}(\mathcal{C}_1) \oplus \cdots \oplus \operatorname{gen}(\mathcal{C}_{p-1}) \oplus \operatorname{gen}(\mathcal{C}_p).$$

Por lo tanto, $\beta = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_{p-1} \cup \mathcal{C}_p$ es una base de $\mathcal{U}_p$.

Finalmente, para $i \in \{1, \ldots, t_p\}$ definimos

$$\mathcal{G}_p^{(i)} = \operatorname{gen}\{\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i)}, \ldots, \mathbf{x}_p^{(i)}\}$$

y para $j \in \{1, 2, \ldots, p-1\}$ e $i \in \{t_p + \cdots + t_{j+1} + 1, \ldots, t_p + \cdots + t_j\}$, definimos

$$\mathcal{G}_j^{(i)} = \operatorname{gen}\{\mathbf{x}_1^{(i)}, \mathbf{x}_2^{(i)}, \ldots, \mathbf{x}_j^{(i)}\}.$$

Los subespacios $\mathcal{G}_j^{(i)}$ son los generados por los vectores de las filas del diagrama (3.3). Por construcción, cada fila en (3.3) es una cadena de Jordan, así que los subespacios $\mathcal{G}_j^{(i)}$ son subespacios de Jordan asociados a λ. Además, $\dim \mathcal{G}_j^{(i)} = j$; y como el orden de los vectores de una base no afecta independencia lineal ni el generado, por aplicación reiterada del Teorema 1.3.6,

$$\mathcal{U}_p = (\mathcal{G}_p^{(1)} \oplus \cdots \oplus \mathcal{G}_p^{(t_p)}) \oplus \cdots \oplus (\mathcal{G}_1^{(t_p+\cdots+t_2+1)} \oplus \cdots \oplus \mathcal{G}_1^{(t_p+\cdots+t_2+t_1)}).$$

□

3.3. Caracterización del índice de un valor propio

En esta sección mostramos un resultado que establece que el índice de un valor propio λ de una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es igual a la multiplicidad de λ como raíz del polinomio minimal de $\mathbf{A}$. Este resultado se obtiene como aplicación del Teorema 3.2.1.

Teorema 3.3.1. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\operatorname{Spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_k\}$ y $m_{\mathbf{A}}(x) = \prod_{i=1}^{k} (x - \lambda_i)^{m_i}$. Entonces

$$i_{\mathbf{A}}(\lambda_j) = m_j,$$

para todo $j \in \{1, \ldots, k\}$.

Demostración. Sea $j \in \{1, \ldots, k\}$. Tenemos que demostrar que

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1}) \subsetneqq \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}) = \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j + 1}).$$

De esto se seguirá que $i_{\mathbf{A}}(\lambda_j) = m_j$. Veamos inicialmente que

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1}) \subsetneqq \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}).$$

Razonemos por reducción al absurdo. Supongamos que

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1}) = \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}).$$

Por el Teorema 3.2.1

$$\mathbb{C}^n = \sum_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{m_i}).$$

Sea $q(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_j)^{m_j - 1} \cdots (x - \lambda_k)^{m_k}$. Entonces $q(x)$ es un polinomio mónico no nulo de grado menor que $m_{\mathbf{A}}(x)$. Luego, cada $\mathbf{x} \in \mathbb{C}^n$ puede expresarse en la forma

$$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_j + \cdots + \mathbf{x}_k.$$

donde $\mathbf{x}_i \in \mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{m_i})$ para cada $i \in \{1, \ldots, k\}$. Notemos que, para $i = j$ también tenemos que $((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1})\mathbf{x}_j = 0$ por nuestra suposición.

Definamos $t_i = m_i$ para todo $i \in \{1, \ldots, k\} - \{j\}$ y $t_j = m_j - 1$. Luego, conmutando adecuadamente tenemos que

$$\begin{aligned}
q(\mathbf{A})\mathbf{x} &= (\mathbf{A} - \lambda_1 \mathbf{I}_n)^{m_1} \cdots (\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1} \cdots (\mathbf{A} - \lambda_k \mathbf{I}_n)^{m_k} \left(\sum_{i=1}^{k} \mathbf{x}_i \right) \\
&= \sum_{i=1}^{k} \left[(\mathbf{A} - \lambda_1 \mathbf{I}_n)^{m_1} \cdots (\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1} \cdots (\mathbf{A} - \lambda_k \mathbf{I}_n)^{m_k} \mathbf{x}_i \right] \\
&= \sum_{i=1}^{k} \left[\left(\prod_{\substack{s=1 \\ s \neq i}}^{k} (\mathbf{A} - \lambda_s \mathbf{I}_n)^{t_s} \right) (\mathbf{A} - \lambda_i \mathbf{I}_n)^{t_i} \mathbf{x}_i \right] \\
&= \sum_{i=1}^{k} \left(\prod_{\substack{s=1 \\ s \neq i}}^{k} (\mathbf{A} - \lambda_s \mathbf{I}_n)^{t_s} \right) 0 \\
&= \sum_{i=1}^{k} 0 \\
&= 0.
\end{aligned}$$

Por Lema 1.3.10, $q(\mathbf{A}) = 0$ y por lo tanto $q(x) \in Ann(\mathbf{A})$, lo cual contradice la minimalidad del grado de $m_{\mathbf{A}}(x)$. La contradicción proviene de haber supuesto que $\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1}) = \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j})$. Hemos probado así que

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j - 1}) \subsetneqq \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}).$$

Ahora veamos que $\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}) = \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j + 1})$.

De nuevo, por reducción al absurdo supongamos que

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}) \subsetneqq \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j + 1}).$$

Por el Teorema 3.2.1 tenemos que

$$\mathbb{C}^n = \sum_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{m_i}),$$

y como

$$\mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j}) \subsetneqq \mathcal{N}((\mathbf{A} - \lambda_j \mathbf{I}_n)^{m_j + 1}),$$

también se cumple que

$$\mathbb{C}^n = \sum_{i=1}^{k} \mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{s_i}),$$

donde $s_i = m_i$ para todo $i \in \{1, \ldots, k\} - \{j\}$ y $s_j = m_j + 1$. Ahora, sean $u \in \{1, \ldots, k\}$ y

$$\mathbf{x} \in \mathcal{N}((\mathbf{A} - \lambda_u \mathbf{I}_n)^{s_u}) \bigcap \sum_{\substack{v=1 \\ v \neq u}}^{k} \mathcal{N}((\mathbf{A} - \lambda_v \mathbf{I}_n)^{s_v}).$$

Entonces $((\mathbf{A} - \lambda_u \mathbf{I}_n)^{s_u})\mathbf{x} = 0$ y $\mathbf{x} = \sum\limits_{r=1, r \neq u}^{k} \mathbf{x}_r$, donde $\mathbf{x}_r \in \mathcal{N}((\mathbf{A} - \lambda_r \mathbf{I}_n)^{s_r})$ para todo $r \in \{1, \ldots, k\} - \{u\}$. Ahora, como $(x - \lambda_u)^{s_u}$ y $\prod\limits_{v=1, v \neq u}^{k} (x - \lambda_v)^{s_v}$ son primos relativos, existen $s(x), t(x) \in \mathbb{C}\,[x]$ tales que

$$1 = s(x)(x - \lambda_u)^{s_u} + t(x) \prod_{\substack{v=1 \\ v \neq u}}^{k} (x - \lambda_v)^{s_v}.$$

Luego tenemos

$$\begin{aligned}
\mathbf{x} &= \mathbf{I}_n\mathbf{x} \\
&= \left(s(\mathbf{A})(\mathbf{A}-\lambda_u\mathbf{I}_n)^{s_u} + t(\mathbf{A}) \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v} \right) \mathbf{x} \\
&= s(\mathbf{A})(\mathbf{A}-\lambda_u\mathbf{I}_n)^{s_u}\mathbf{x} + t(\mathbf{A}) \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}\mathbf{x} \\
&= s(\mathbf{A})(0) + t(\mathbf{A}) \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}\mathbf{x} \\
&= t(\mathbf{A}) \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v} \left(\sum_{\substack{r=1\\r\neq u}}^{k} \mathbf{x}_r \right) \\
&= t(\mathbf{A}) \sum_{\substack{r=1\\r\neq u}}^{k} \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}\mathbf{x}_r.
\end{aligned}$$

Por otro lado tenemos que para todo $r \in \{1,\ldots,k\} - \{u\}$

$$\begin{aligned}
\prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}\mathbf{x}_r &= \left(\prod_{\substack{v=1\\v\neq u,r}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v} \right) (\mathbf{A}-\lambda_r\mathbf{I}_n)^{s_r}\mathbf{x}_r \\
&= \left(\prod_{\substack{v=1\\v\neq u,r}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v} \right) (0) \\
&= 0.
\end{aligned}$$

Así,

$$\begin{aligned}
\mathbf{x} &= t(\mathbf{A}) \sum_{\substack{r=1\\r\neq u}}^{k} \prod_{\substack{v=1\\v\neq u}}^{k} (\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}\mathbf{x}_r \\
&= t(\mathbf{A}) \left(\sum_{\substack{r=1\\r\neq u}}^{k} 0 \right) \\
&= t(\mathbf{A})(0) \\
&= 0.
\end{aligned}$$

Por lo tanto, $\mathcal{N}((\mathbf{A}-\lambda_u\mathbf{I}_n)^{s_u}) \bigcap \sum_{v=1, v\neq u}^{k} \mathcal{N}((\mathbf{A}-\lambda_v\mathbf{I}_n)^{s_v}) = \{0\}$. En consecuencia,

$$\mathbb{C}^n = \bigoplus_{i=1}^{k} \mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{s_i}).$$

Pero esto nos genera una contradicción, puesto que tenemos ahora

$$\begin{aligned} n = \dim(\mathbb{C}^n) &= \dim\left(\bigoplus_{i=1}^{k} \mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{s_i})\right) \\ &= \sum_{i=1}^{k} \dim(\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{s_i})) \\ &> \sum_{i=1}^{k} \dim(\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i})) \\ &= n. \end{aligned}$$

La contradicción proviene de haber supuesto que

$$\mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}) \subsetneq \mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j+1}),$$

así que debemos tener realmente $\mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j}) = \mathcal{N}((\mathbf{A}-\lambda_j\mathbf{I}_n)^{m_j+1})$. Esto termina la demostración. $\square$

Capítulo 4

Teorema de Jordan y aplicaciones

En este capítulo realizamos una demostración de nuestro teorema principal, el teorema de Jordan. Gran parte de la prueba está contenida en los resultados del Capítulo 3. Primero introducimos los conceptos de bloques y matrices de Jordan que nos servirán para formular el teorema de Jordan. Veremos algunas propiedades de este tipo de matrices, y finalmente mostraremos algunas consecuencias del teoreoma de Jordan.

4.1. El teorema de Jordan

Definición 4.1.1. Un **bloque de Jordan**, $\mathbf{J}_k(\lambda)$, es una matriz triangular superior de $\mathbb{C}^{k\times k}$ de la forma:

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & & 1 \\ 0 & \cdots & \cdots & & \lambda \end{bmatrix} = \begin{bmatrix} \lambda\mathbf{e}_1 & \mathbf{e}_1 + \lambda\mathbf{e}_2 & \cdots & \mathbf{e}_{k-1} + \lambda\mathbf{e}_k \end{bmatrix}$$

Una **matriz de Jordan** en $\mathbb{C}^{n\times n}$ es una matriz diagonal por bloques de la forma

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{n_1}(\lambda_1) & & & \\ & \mathbf{J}_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & \mathbf{J}_{n_s}(\lambda_s) \end{bmatrix} \tag{4.1}$$

donde $n_1 + n_2 + \cdots + n_s = n$ y los λ_i son números complejos no necesariamente distintos.

El teorema de Jordan afirma que toda matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es semejante a una matriz de Jordan. Antes de demostrar el teorema de Jordan de manera general, demostramos el caso particular en el que la matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ tiene un solo valor propio.

Lema 4.1.1. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$ tal que $\mathrm{Spec}(\mathbf{A}) = \{\lambda\}$. Entonces $\mathbf{A}$ es semejante a una matriz de Jordan con bloques de la forma $\mathbf{J}_k(\lambda)$.

Demostración. Como $\mathrm{Spec}(\mathbf{A}) = \{\lambda\}$, tenemos que $m_{\mathbf{A}}(x) = (x-\lambda)^m$, donde $1 \leq m \leq n$. Por el Teorema 3.2.1

$$\mathbb{C}^n = \mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^m),$$

y por el Teorema 3.2.3, $\mathcal{N}((\mathbf{A} - \lambda\mathbf{I}_n)^m)$ es suma directa de subespacios de Jordan asociados a λ. Luego, $\mathbb{C}^n$ es suma directa de subespacios cíclicos de Jordan asociados a λ. Por lo tanto, existen cadenas de Jordan asociadas a λ

$$\begin{gathered} \mathbf{x}_{k_1}^{(1)}, \ldots, \mathbf{x}_1^{(1)} \\ \mathbf{x}_{k_2}^{(2)}, \ldots, \mathbf{x}_1^{(2)} \\ \vdots \\ \mathbf{x}_{k_s}^{(s)}, \ldots, \mathbf{x}_1^{(s)} \end{gathered}$$

donde $\left\{\mathbf{x}_1^{(1)}, \mathbf{x}_1^{(2)}, \ldots, \mathbf{x}_1^{(s)}\right\}$ es una base de $E_{\mathbf{A}}(\lambda)$, $k_1 + \cdots + k_s = n$, y la unión de las estas cadenas de Jordan forma una base de $\mathbb{C}^n$. Así, por Lema 1.3.7, la matriz

$$\mathbf{P} = \begin{bmatrix} \mathbf{x}_1^{(1)} & \cdots & \mathbf{x}_{k_1}^{(1)} & \cdots & \mathbf{x}_1^{(s)} & \cdots & \mathbf{x}_{k_s}^{(s)} \end{bmatrix} \in \mathbb{C}^{n\times n}$$

es invertible. Tenemos que, $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ es igual a

$$
\begin{aligned}
&\mathbf{P}^{-1}\mathbf{A}\begin{bmatrix}\mathbf{x}_1^{(1)} & \cdots & \mathbf{x}_{k_1}^{(1)} & \cdots & \mathbf{x}_1^{(s)} & \cdots & \mathbf{x}_{k_s}^{(s)}\end{bmatrix}\\
&=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{A}\mathbf{x}_1^{(1)} & \cdots & \mathbf{A}\mathbf{x}_{k_1}^{(1)} & \cdots & \mathbf{A}\mathbf{x}_1^{(s)} & \cdots & \mathbf{A}\mathbf{x}_{k_s}^{(s)}\end{bmatrix}\\
&=\mathbf{P}^{-1}\begin{bmatrix}\lambda\mathbf{x}_1^{(1)} & \cdots & \lambda\mathbf{x}_{k_1}^{(1)}+\mathbf{x}_{k_1-1}^{(1)} & \cdots & \lambda\mathbf{x}_1^{(s)} & \cdots & \lambda\mathbf{x}_{k_s}^{(s)}+\mathbf{x}_{k_s-1}^{(s)}\end{bmatrix}\\
&=\begin{bmatrix}\lambda\mathbf{P}^{-1}\mathbf{x}_1^{(1)} & \cdots & \mathbf{P}^{-1}(\lambda\mathbf{x}_{k_1}^{(1)}+\mathbf{x}_{k_1-1}^{(1)}) & \cdots & \lambda\mathbf{P}^{-1}\mathbf{x}_1^{(s)} & \cdots & \mathbf{P}^{-1}(\lambda\mathbf{x}_{k_s}^{(s)}+\mathbf{x}_{k_s-1}^{(s)})\end{bmatrix}\\
&=\begin{bmatrix}\lambda\mathbf{e}_1 & \cdots & \lambda\mathbf{e}_{k_1}+\mathbf{e}_{(k_1-1)} & \cdots & \lambda\mathbf{e}_{(k_1+\cdots+k_{s-1}+1)} & \cdots & \lambda\mathbf{e}_{(k_1+\cdots+k_s)}+\mathbf{e}_{(k_1+\cdots+k_s-1)}\end{bmatrix}\\
&=\begin{bmatrix}\mathbf{J}_{k_1}(\lambda) & & \\ & \ddots & \\ & & \mathbf{J}_{k_s}(\lambda)\end{bmatrix}.
\end{aligned}
$$

□

Ahora estamos en posición de demostrar nuestro teorema principal.

Teorema 4.1.2 (Teorema de Jordan)**.** Toda matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es semejante a una matriz de Jordan.

Demostración. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\operatorname{Spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_k\}$ y para cada $i \in \{1, 2, \ldots, k\}$, sea m_i la multiplicidad de λ_i como raíz de $m_{\mathbf{A}}(x)$. Entonces

$$m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{m_i}$$

y aplicando el Teorema 3.2.1 obtenemos la descomposición

$$\mathbb{C}^n = \bigoplus_{i=1}^{k}\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i}).$$

Por el Teorema 3.2.3, $\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i})$ es suma directa de subespacios cíclicos de Jordan asociados a λ_i, $i = 1, \ldots, k$. Entonces existe una base β de $\mathbb{C}^n$ tal que

$$\beta = \beta_1 \cup \cdots \cup \beta_k,$$

donde β_i es base de $\mathcal{N}((\mathbf{A}-\lambda_i\mathbf{I}_n)^{m_i})$, para todo $i \in \{1, \ldots, k\}$. Además, para cada $i \in \{1, 2, \ldots, k\}$, β_i es unión de cadenas de Jordan asociadas a λ_i.

Para cada $i \in \{1, 2, \ldots, k\}$, organizamos β_i como en la demostración del Lema 4.1.1, y organizamos β de tal manera que las cadenas de Jordan asociadas a λ_1 aparezcan consecutivas, y que sigan las cadenas de Jordan asociadas a λ_2 en forma consecutiva, y así sucesivamente hasta λ_k.

Definimos una matriz $\mathbf{P} \in \mathbb{C}^n$ de tal manera que sus columnas sean los vectores de β en el orden que hemos descrito. Entonces, por el Lema 1.3.7, $\mathbf{P}$ es invertible y por analogía a la demostración del Lema 4.1.1 resulta que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{J}_1 & & \\ & \ddots & \\ & & \mathbf{J}_k \end{bmatrix},$$

donde

$$\mathbf{J}_i = \begin{bmatrix} \mathbf{J}_{k_{i1}}(\lambda_i) & & \\ & \ddots & \\ & & \mathbf{J}_{k_{is_i}}(\lambda_i) \end{bmatrix},$$

para todo $i \in \{1, \ldots, k\}$, y k_{ij} es la longitud de la $j-$ésima cadena de Jordan asociada a λ_i, $j = 1, \ldots, s_i$.

Hemos mostrado que $\mathbf{A}$ es semejante a una matriz de Jordan, lo que termina la demostración. $\square$

Observemos que en la parte final de la demostración anterior, se satisface que

$$\sum_{i=1}^{k} \sum_{j=1}^{s_i} k_{ij} = n.$$

Para una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$, una matriz de Jordan $\mathbf{J}$ semejante a $\mathbf{A}$ es llamada un **forma canónica de Jordan para A**.

A continuación analizamos con un poco más de profundidad la demostración del Teorema 4.1.2, y obtendremos algunas observaciones importantes acerca de las matrices de Jordan semejantes a $\mathbf{A} \in \mathbb{C}^{n\times n}$.

En la demostración del Teorema 4.1.2 notamos que por cada cadena de Jordan de longitud r asociada a algún valor propio λ de $\mathbf{A}$, la matriz de Jordan $\mathbf{J}$ tiene un bloque de Jordan del tipo $\mathbf{J}_r(\lambda)$.

También notamos que la manera de ordenar las cadenas de Jordan presentes en la base β para armar la matriz $\mathbf{P}$ podría haberse hecho de diferentes formas, no necesariamente poniendo primero las asociadas a λ_1, luego las asociadas a λ_2, y así sucesivamente, sino que podemos poner primero cualquiera de las cadenas de Jordan asociadas a algún λ_{j_1}, luego una cadena de Jordan asociada a algún λ_{j_2} (no necesariamente distinto de λ_{j_1}, y tampoco tiene que ser igual a λ_{j_1}), luego una tercera cadena de Jordan asociada a algún λ_{j_3} (no necesariamente distinto de λ_{j_1} y λ_{j_2}), y así sucesivamente hasta que hallamos agotado todas las cadenas de Jordan contenidas en β. La única condición es que los vectores en cada cadena de Jordan contenida en β aparezcan de manera consecutiva en la forma $\mathbf{x}_r, \ldots, \mathbf{x}_1$, donde

$$x_{r-k} = (\lambda \mathbf{I}_n - \mathbf{A})^k \mathbf{x}_r, \quad k = 0, \ldots, r-1,$$

y λ es algún valor propio de $\mathbf{A}$. Los vectores de esta cadena de Jordan aparecerán como columnas consecutivas en la matriz $\mathbf{P}$, y la matriz $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ tendrá un bloque de Jordan de la forma $\mathbf{J}_r(\lambda)$ que corresponde a la cadena de Jordan $\mathbf{x}_r, \ldots, \mathbf{x}_1$. Así, la matriz $\mathbf{J}$ es semejante a $\mathbf{A}$, pero el orden de los bloques de Jordan presentes en $\mathbf{J}$ se corresponde con el orden escogido para las cadenas de Jordan en β.

Luego, podemos afirmar que si dos matrices de Jordan $\mathbf{J}_1$ y $\mathbf{J}_2$ semejantes a $\mathbf{A}$ se construyen de esta forma, entonces las matrices $\mathbf{J}_1$ y $\mathbf{J}_2$ tienen los mismos bloques de Jordan, pero ordenados de manera posiblemente diferente.

Ahora mostraremos que cualquier matriz de Jordan $\mathbf{J}$ semejante a $\mathbf{A}$ se obtiene de esta manera, es decir, que $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ donde las columnas de $\mathbf{P}$ forman una base de $\mathbb{C}^n$, la cual es unión de cadenas de Jordan asociadas a algún valor propio de $\mathbf{A}$, y que en $\mathbf{P}$, las columnas están ordenadas de tal manera que los vectores de cada cadena de Jordan aparecen de manera consecutiva. En efecto, tenemos que existe una matriz $\mathbf{P}$ invertible tal que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$. Sea $\mathbf{J}_k(\lambda)$ un bloque de Jordan de $\mathbf{J}$ y digamos que el bloque $\mathbf{J}_k(\lambda)$ aparece desde la columna j_1 hasta la columna $j_2 = j_1 + k - 1$. Esto significa que las columnas de $\mathbf{J}$ que corresponden al bloque $\mathbf{J}_k(\lambda)$ son

$$\lambda \mathbf{e}_{j_1}, \lambda \mathbf{e}_{j_1+1} + \mathbf{e}_{j_1}, \ldots, \lambda \mathbf{e}_{j_2} + \mathbf{e}_{j_2-1}.$$

Ahora, teniendo en cuenta que $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$, vemos que las columnas de $\mathbf{J}$ anteriores se obtienen multiplicando por $\mathbf{P}^{-1}\mathbf{A}$ las columnas correspondientes de la matriz $\mathbf{P}$, a saber,

$$\mathbf{P}_{*j_1}, \ldots, \mathbf{P}_{*j_2},$$

lo que significa que

$$\begin{aligned}
\mathbf{P}^{-1}\mathbf{A}\mathbf{P}_{*j_1} &= \lambda \mathbf{e}_{j_1} \\
\mathbf{P}^{-1}\mathbf{A}\mathbf{P}_{*(j_1+1)} &= \lambda \mathbf{e}_{j_1+1} + \mathbf{e}_{j_1} \\
&\vdots \\
\mathbf{P}^{-1}\mathbf{A}\mathbf{P}_{*j_2} &= \lambda \mathbf{e}_{j_2} + \mathbf{e}_{j_2-1}
\end{aligned}$$

y luego

$$\begin{aligned}
\mathbf{A}\mathbf{P}_{*j_1} &= \mathbf{P}(\lambda \mathbf{e}_{j_1}) = \lambda \mathbf{P}_{*j_1} \\
\mathbf{A}\mathbf{P}_{*(j_1+1)} &= \mathbf{P}(\lambda \mathbf{e}_{j_1+1} + \lambda \mathbf{e}_{j_1}) = \lambda \mathbf{P}_{*(j_1+1)} + \mathbf{P}_{*j_1} \\
&\vdots \\
\mathbf{A}\mathbf{P}_{*j_2} &= \mathbf{P}(\lambda \mathbf{e}_{j_2} + \mathbf{e}_{j_2-1}) = \lambda \mathbf{P}_{*j_2} + \mathbf{P}_{*(j_2-1)}.
\end{aligned}$$

Estas ecuaciones significan precisamente que $\mathbf{P}_{*j_2}, \mathbf{P}_{*(j_2-1)} \ldots, \mathbf{P}_{*j_1}$ es una cadena de Jordan asociada a λ de longitud k. Esto muestra que en la matriz $\mathbf{P}$, las columnas son cadenas de Jordan sucesivas asociadas a algún valor propio de $\mathbf{A}$. Además, como $\mathbf{P}$ es invertible, concluimos que las columnas de $\mathbf{P}$ forma una base de $\mathbb{C}^n$ la cual es una unión de cadenas de Jordan.

4.2. Consecuencias del Teorema de Jordan

En esta última sección mostramos algunas consecuencias del teorema de Jordan, así como propiedades de las matrices de Jordan.

Teorema 4.2.1. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, λ un valor propio de $\mathbf{A}$ de índice p, y $\mathbf{J} \in \mathbb{C}^{n\times n}$ una matriz de Jordan semejante a $\mathbf{A}$. Entonces el número de bloques de Jordan en $\mathbf{J}$ asociados a λ es igual a $\mathrm{mg}_{\mathbf{A}}(\lambda)$ y

$$\dim(\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p)) = \mathrm{ma}_{\mathbf{A}}(\lambda).$$

Demostración. Como hemos visto, puede asumirse que $\mathbf{J}$ se obtiene como en la demostración del Teorema 4.1.2, pues de los contrario, la única diferencia está en el orden de los bloques de Jordan que aparecen en $\mathbf{J}$. En todo caso, la cantidad de bloques de Jordan en $\mathbf{J}$ es la misma.

En la demostración del Teorema 3.2.3 se construye una base β de $\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p)$, ver diagrama (3.3). En (3.3) los vectores están organizados de tal forma que las filas son cadenas de Jordan asociadas a λ y cada una de esas filas contiene un único valor propio de $\mathbf{A}$. Además, el conjunto de tales vectores propios de $\mathbf{A}$ (primera columna de (3.3)) es una base para $E_{\mathbf{A}}(\lambda)$. Por lo tanto, la dimensión de $E_{\mathbf{A}}(\lambda)$, es decir $\mathrm{mg}_{\mathbf{A}}(\lambda)$, es igual al número de filas en (3.3), pero por cada fila de esas hay exactamente un bloque de Jordan en $\mathbf{J}$ asociado a λ. Esto muestra la primera parte del teorema.

Ahora, el número de veces que aparece λ en la diagonal principal de $\mathbf{J}$ es $\mathrm{ma}_{\mathbf{A}}(\lambda)$, y este número también es igual a la sumatoria de los tamaños de todos los bloques de Jordan en $\mathbf{J}$ asociados a λ, que a su vez es igual a la sumatoria de las longitudes de las cadenas de Jordan asociadas a λ que forman la base de $\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p)$. Por lo tanto

$$\dim(\mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^p)) = \mathrm{ma}_{\mathbf{A}}(\lambda).$$

□

Teorema 4.2.2. Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$. Entonces $m_{\mathbf{A}}(x) = P_{\mathbf{A}}(x)$ si y solo si para todo $\lambda \in \mathrm{Spec}(\mathbf{A})$ se cumple que $\mathrm{mg}_{\mathbf{A}}(\lambda) = 1$.

Demostración. Sean $\mathbf{J}$ una matriz de Jordan semejante a $\mathbf{A}$ y $\mathrm{Spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_k\}$. Entonces podemos escribir

$$m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{m_i}, \quad P_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{n_i}$$

donde $0 < m_i \leq n_i$, $i = 1, \ldots, k$. Así que $m_{\mathbf{A}}(x) = P_{\mathbf{A}}(x)$ si y solo si para todo $i \in \{1, \ldots, k\}$ se cumple que $m_i = n_i$.

Ahora, por el Teorema 3.3.1, para cada $i \in \{1, \dots, k\}$ se tiene que $i_{\mathbf{A}}(\lambda_i) = m_i$; y por el Teorema 4.2.1, n_i es la dimensión de $\mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{i_{\mathbf{A}}(\lambda_i)})$.

Por lo tanto tenemos que $m_{\mathbf{A}}(x) = P_{\mathbf{A}}(x)$ si y solo

$$i_{\mathbf{A}}(\lambda_i) = \dim(\mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{i_{\mathbf{A}}(\lambda_i)}))$$

para $i = 1, \dots, k$. Con referencia al diagrama (3.3), notamos que $i_{\mathbf{A}}(\lambda_i)$ es la cantidad de columnas, mientras $\dim(\mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{i_{\mathbf{A}}(\lambda_i)}))$ es la cantidad total de vectores presentes en dicho diagrama. Y notamos que estas cantidades son iguales si y solamente si en (3.3) hay una sola fila.

Ahora, la multiplicidad geométrica de λ_i es igual a la cantidad de filas en el respectivo diagrama (3.3), así que si $m_{\mathbf{A}}(x) = P_{\mathbf{A}}(x)$, entonces $\mathrm{mg}_{\mathbf{A}}(\lambda_i) = 1$ para todo $i \in \{1, \dots, k\}$.

Recíprocamente, si $\mathrm{mg}_{\mathbf{A}}(\lambda_i) = 1$ para todo $i \in \{1, \dots, k\}$, entonces, para cada $i \in \{1, \dots, k\}$, en el respectivo diagrama (3.3) se tiene una sola fila; esta fila es una cadena de Jordan de longitud $i_{\mathbf{A}}(\lambda_i) = m_i$, y por otro lado la cantidad total de vectores en esta única fila es igual a $\mathcal{N}((\mathbf{A} - \lambda_i \mathbf{I}_n)^{i_{\mathbf{A}}(\lambda_i)}) = n_i$. Concluimos que $m_i = n_i$ para todo $i \in \{1, \dots, k\}$ y por lo tanto, $m_{\mathbf{A}}(x) = P_{\mathbf{A}}(x)$. □

Teorema 4.2.3. Sea $\mathbf{A} \in \mathbb{C}^{n \times n}$. Entonces $\mathbf{A}$ es semejante a $\mathbf{A}^T$.

Demostración. Por el teorema de Jordan, $\mathbf{A}$ es semejante a una matriz de Jordan

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{n_1} & & \\ & \ddots & \\ & & \mathbf{J}_{n_k} \end{bmatrix},$$

donde $\mathbf{J}_{n_i}$ es un bloque de Jordan de tamaño $n_i \times n_i$, para $i = 1, \dots, k$ y $n_1 + \cdots + n_k = n$. Para $i = 1, \dots, k$, sea

$$\mathbf{T}_i = \begin{bmatrix} \mathbf{e}_{n_i} & \mathbf{e}_{n_i - 1} & \cdots & \mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}.$$

Observemos que

$$\mathbf{T}_i^2 = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = I_{n_i},$$

de donde $(\mathbf{T}_i)^{-1} = \mathbf{T}_i$. Ahora, para $i = 1, \ldots, k$ tenemos

$$\begin{aligned} (\mathbf{T}_i)^{-1}\mathbf{J}_{n_i}\mathbf{T}_i &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 1 & \lambda \\ 0 & 0 & \cdots & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \lambda & \cdots & 0 & 0 \\ \lambda & 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{bmatrix} \\ &= (\mathbf{J}_{n_i})^T \end{aligned}$$

Definamos la matriz $\mathbf{T}$ diagonal por bloques, en forma similar a una matriz de Jordan, como sigue

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{n_1} & & \\ & \ddots & \\ & & \mathbf{T}_{n_k} \end{bmatrix} \in \mathbb{C}^{n \times n}.$$

Entonces $\mathbf{T} = \mathbf{T}^{-1}$ y $\mathbf{T}^{-1}\mathbf{J}\mathbf{T} = \mathbf{J}^T$. Por lo tanto $\mathbf{J}$ es semejante a $\mathbf{J}^T$. Como $\mathbf{A}$ es semejante a $\mathbf{J}$, entonces existe $\mathbf{P} \in \mathbb{C}^{n\times n}$ invertible tal que $\mathbf{J} = \mathbf{PAP}^{-1}$. Luego,

$$\begin{aligned}
\mathbf{A}^T &= (\mathbf{PJP}^{-1})^T \\
&= (\mathbf{P}^{-1})^T \mathbf{J}^T \mathbf{P}^T \\
&= (\mathbf{P}^T)^{-1} \mathbf{T}^{-1} \mathbf{JTP}^T \\
&= (\mathbf{P}^T)^{-1} \mathbf{T}^{-1} \mathbf{PAP}^{-1} \mathbf{TP}^T \\
&= (\mathbf{P}^{-1}\mathbf{TP}^T)^{-1} \mathbf{A} (\mathbf{P}^{-1}\mathbf{TP}^T).
\end{aligned}$$

Por lo tanto $\mathbf{A}$ es semejante a $\mathbf{A}^T$. □

Mostramos a continuación una caracterización de las matrices diagonalizables en términos de las multiplicidades algebraica y geométrica de sus valores propios.

Teorema 4.2.4. Una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ es diagonalizable si y solo si para todo valor propio λ de $\mathbf{A}$ se tiene

$$\mathrm{mg}_{\mathbf{A}}(\lambda) = \mathrm{ma}_{\mathbf{A}}(\lambda).$$

Equivalentemente, $\mathbf{A}$ es diagonalizable si y solo si el polinomio minimal de $\mathbf{A}$ tiene todos sus ceros de multiplicidad algebraica 1.

Demostración. Supongamos que $\mathbf{A} \in \mathbb{C}^{n\times n}$ es diagonalizable. Entonces existe una matriz diagonal $\mathbf{D}$ semejante a $\mathbf{A}$; pero esta matriz $\mathbf{D}$ es una matriz de Jordan en donde todos los bloques de Jordan de $\mathbf{D}$ son de tamaño 1×1. Ahora, sabemos que podemos escribir $\mathbf{D} = \mathbf{P}^{-1}\mathbf{AP}$, donde las columnas de $\mathbf{P}$ forman una base de $\mathbb{C}^n$, la cual es unión de cadenas de Jordan asociadas a algún valor propio de $\mathbf{A}$, y que en $\mathbf{P}$ las columnas están ordenadas de tal manera que los vectores de cada cadena de Jordan aparecen de manera consecutiva. Fijemos un valor propio λ de $\mathbf{A}$. En el respectivo diagrama (3.3) para λ resulta que todas las cadenas de Jordan (filas del diagrama) tienen longitud 1. Eso significa que hay una sola columna en el diagrama, lo que quiere decir que $i_{\mathbf{A}}(\lambda) = 1$. Así, no hay vectores propios generalizados de orden $r > 1$ asociados a λ. Todos los vectores del respectivo diagrama son vectores propios de $\mathbf{A}$. Pero, la cantidad de filas del diagrama es la multiplicidad geométrica de λ y, por el Teorema 4.2.1, la cantidad total de vectores en el diagrama es la multiplicidad algebraica de λ, así que $\mathrm{mg}_{\mathbf{A}}(\lambda) = \mathrm{ma}_{\mathbf{A}}(\lambda)$.

Recíprocamente, suponer que $\mathrm{mg}_{\mathbf{A}}(\lambda) = \mathrm{ma}_{\mathbf{A}}(\lambda)$ para todo valor propio λ de $\mathbf{A}$ significa que para cada λ, en el respectivo diagrama (3.3), la cantidad de filas coincide con la cantidad total de vectores del diagrama, lo que quiere decir que solo puede haber una columna en tal diagrama. Así, todas las cadenas de Jordan asociadas a λ tienen longitud 1, lo que significa que la matriz de Jordan $\mathbf{J}$ semejante a $\mathbf{A}$ que se construye mediante este proceso, tiene solo bloques de Jordan de tamaño 1×1 asociados a λ. Por lo tanto la matriz $\mathbf{J}$ es una matriz diagonal, y esto muestra que $\mathbf{A}$ es diagonalizable.

Esto muestra que $\mathbf{A}$ es diagonalizable si y solo si para todo valor propio λ de $\mathbf{A}$ se tiene $\mathrm{mg}_{\mathbf{A}}(\lambda) = \mathrm{ma}_{\mathbf{A}}(\lambda)$.

Ahora, si el polinomio minimal de $\mathbf{A}$ es $m_{\mathbf{A}}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{m_i}$, donde $\mathrm{Spec}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_k\}$, y sabiendo para cada i, m_i es la cantidad de columnas del diagrama (3.3) respectivo para λ_i, tenemos que si $m_i = 1$ para todo $i \in \{1, \ldots, k\}$, entonces tales diagramas tienen una sola columna y ya vimos que esto lleva a que $\mathbf{A}$ es diagonalizable. Recíprocamente, si $\mathbf{A}$ es diagonalizable, entonces, para todo $i \in \{1, \ldots, k\}$, el diagrama (3.3) respectivo tiene una sola columna y eso quiere decir que $m_i = 1$. $\square$

Los resultados del resto de la sección establecen propiedades acerca del número de bloques de Jordan de tamaño dado que aparecen en la forma canónica de Jordan de una matriz dada $\mathbf{A} \in \mathbb{C}^{n\times n}$.

Los dos resultados que siguen muestran que dos matrices de Jordan son semejantes si y solo si una se obtiene de la otra permutando sus bloques de Jordan.

Teorema 4.2.5. Sean $\mathbf{J}, \mathbf{J}' \in \mathbb{C}^{n\times n}$ matrices de Jordan tales que $\mathbf{J}'$ se obtiene al realizar una permutación de los bloques de Jordan de $\mathbf{J}$. Entonces $\mathbf{J}$ y $\mathbf{J}'$ son semejantes.

Demostración. Sea

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{n_1}(\lambda_1) & & & & & \\ & \ddots & & & & \\ & & \mathbf{J}_{n_i}(\lambda_i) & & & \\ & & & \mathbf{J}_{n_{i+1}}(\lambda_{i+1}) & & \\ & & & & \ddots & \\ & & & & & \mathbf{J}_{n_s}(\lambda_s) \end{bmatrix}$$

una matriz de Jordan. Supongamos que $\mathbf{J}_{n_i}(\lambda_i)$ tiene su primera fila sobre la fila t de $\mathbf{J}$. Sea

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{t-1}\mathbf{e}_{t+n_i} & \cdots & \mathbf{e}_{t+n_i+n_{i+1}-1}\mathbf{e}_t & \cdots & \mathbf{e}_{t+n_i-1}\mathbf{e}_{t+n_i+n_{i+1}} & \cdots & \mathbf{e}_n \end{bmatrix} \in \mathbb{C}^{n\times n},$$

entonces $\mathbf{E}^{-1}\mathbf{J}\mathbf{E}$ es igual a la matriz

$$\begin{bmatrix} \mathbf{J}_{n_1}(\lambda_1) & & & & & & & \\ & \ddots & & & & & & \\ & & \mathbf{J}_{n_{i-1}}(\lambda_{i-1}) & & & & & \\ & & & \mathbf{J}_{n_{i+1}}(\lambda_{i+1}) & & & & \\ & & & & \mathbf{J}_{n_i}(\lambda_i) & & & \\ & & & & & \mathbf{J}_{n_{i+2}}(\lambda_{i+2}) & & \\ & & & & & & \ddots & \\ & & & & & & & \mathbf{J}_{n_s}(\lambda_s) \end{bmatrix}$$

Por lo tanto, $\mathbf{J}$ es semejante a cualquier matriz de Jordan que se obtenga al intercambiarle a $\mathbf{J}$ dos de sus bloques de Jordan consecutivos. Ahora, si $\mathbf{J}'$ se obtiene al realizar una permutación de los bloques de Jordan de $\mathbf{J}$, entonces por aplicación reiterada de lo que acabamos de probar y por la transitividad de la semejanza de matrices tenemos que $\mathbf{J}$ y $\mathbf{J}'$ son semejantes. $\square$

Antes del siguiente teorema, notemos que para una matriz de Jordan

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{n_1}(\lambda_1) & & & \\ & \mathbf{J}_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & \mathbf{J}_{n_k}(\lambda_k) \end{bmatrix}$$

tenemos que para cualquier $k \in \mathbb{Z}^+$

$$\mathbf{J}^k = \begin{bmatrix} (\mathbf{J}_{n_1}(\lambda_1))^k & & & \\ & (\mathbf{J}_{n_2}(\lambda_2))^k & & \\ & & \ddots & \\ & & & (\mathbf{J}_{n_k}(\lambda_k))^k \end{bmatrix}$$

Si en $\mathbf{J}$ hay un bloque de Jordan de la forma

$$\mathbf{J}_r(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}$$

con $\lambda \neq 0$, entonces en todas las entradas de la diagonal principal de $\mathbf{J}_r(\lambda)^k$ aparecerá λ^k, así que $\mathbf{J}_r(\lambda)^k$ tiene todas sus columnas no nulas.

Por otro lado, si $\lambda = 0$ tenemos

$$(\mathbf{J}_r(0))^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix};$$

$$(\mathbf{J}_r(0))^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix};$$

y así sucesivamente. Vemos que un bloque de la forma $\mathbf{J}_r(0)$ tiene la primera columna nula, $(\mathbf{J}_r(0))^2$ tiene las dos primeras columnas nulas, $(\mathbf{J}_r(0))^3$ las tres primeras columnas nulas, y así sucesivamente, hasta que la $r-$ésima potencia de $\mathbf{J}_r(0)$ tiene todas sus columnas nulas.

Teorema 4.2.6. Sean $\mathbf{J}, \mathbf{J}' \in \mathbb{C}^{n\times n}$ matrices de Jordan semejantes. Entonces $\mathbf{J}'$ se obtiene al realizar una permutación de los bloques de Jordan de $\mathbf{J}$.

Demostración. Supongamos que $\operatorname{Spec}(\mathbf{J}) = \{\lambda_1, \ldots, \lambda_k\}$. Como $\mathbf{J}$ y $\mathbf{J}'$ son semejantes, entonces por el Teorema 1.5.5

$$P_{\mathbf{J}}(x) = P_{\mathbf{J}'}(x),\ \operatorname{Spec}(\mathbf{J}') = \{\lambda_1, \ldots, \lambda_k\}\ \text{ y } \operatorname{ma}_{\mathbf{J}}(\lambda_i) = \operatorname{ma}_{\mathbf{J}'}(\lambda_i)$$

para todo $i \in \{1, \ldots, k\}$. Para todo $i \in \{1, \ldots, k\}$, sea $n_i = \operatorname{ma}_{\mathbf{J}}(\lambda_i) = \operatorname{ma}_{\mathbf{J}'}(\lambda_i)$. Entonces, en las diagonales principales de $\mathbf{J}$ y $\mathbf{J}'$, el valor propio λ_i aparece exactamente n_i veces.

Como $\mathbf{J}$ y $\mathbf{J}'$ son semejantes, $\mathbf{J} - \lambda_i\mathbf{I}_n$ y $\mathbf{J}' - \lambda_i\mathbf{I}_n$ también lo son. Luego, los bloques de la forma $\mathbf{J}_s(\lambda_i)$ en $\mathbf{J}$ (respectivamente $\mathbf{J}'$) se corresponden en tamaños y cantidades a los bloques de la forma $\mathbf{J}_s(0)$ en $\mathbf{J} - \lambda_i\mathbf{I}_n$ (respectivamente $\mathbf{J}' - \lambda_i\mathbf{I}_n$). La matriz $\mathbf{J} - \lambda_i\mathbf{I}_n$ tiene b_i columnas nulas, donde b_i es el número de bloques de la forma $\mathbf{J}_s(0)$ en $\mathbf{J} - \lambda_i\mathbf{I}_n$ (cada columna de $\mathbf{J} - \lambda_i\mathbf{I}_n$ que contiene a la primera columna de cada bloque de la forma $\mathbf{J}_s(0)$ es nula). Ahora,

$$\mathcal{N}(\mathbf{J} - \lambda_i\mathbf{I}_n) = E_{\mathbf{J}}(\lambda_i),$$

y sabemos que por cada bloque de Jordan en $\mathbf{J}$ de la forma $\mathbf{J}_s(\lambda_i)$, hay un único vector propio que sirve para armar una base de $E_{\mathbf{J}}(\lambda_i)$. Por lo tanto, la dimensión de $E_{\mathbf{J}}(\lambda_i)$ también es igual a b_i, esto es, $\dim(\mathcal{N}(\mathbf{J} - \lambda_i\mathbf{I}_n)) = b_i$. Ahora, la ecuación

$$\dim(\mathcal{N}(\mathbf{J} - \lambda_i\mathbf{I}_n)) + \operatorname{rk}(\mathbf{J} - \lambda_i\mathbf{I}_n) = n$$

muestra que

$$\operatorname{rk}(\mathbf{J} - \lambda_i\mathbf{I}_n) = n - b_i.$$

Análogamente, si $\mathbf{J}' - \lambda_i\mathbf{I}_n$ tiene b_i' bloques de la forma $\mathbf{J}_s(0)$, entonces

$$\operatorname{rk}(\mathbf{J}' - \lambda_i\mathbf{I}_n) = n - b_i'.$$

Como $\mathbf{J} - \lambda_i \mathbf{I}_n$ es semejante a $\mathbf{J}' - \lambda_i \mathbf{I}_n$, tenemos que

$$n - b_i = \mathrm{rk}(\mathbf{J} - \lambda_i \mathbf{I}_n) = \mathrm{rk}(\mathbf{J}' - \lambda_i \mathbf{I}_n) = n - b_i'.$$

Por lo tanto $b_i = b_i'$, es decir, $\mathbf{J} - \lambda_i \mathbf{I}_n$ y $\mathbf{J}' - \lambda_i \mathbf{I}_n$ tienen la misma cantidad de bloques de la forma $\mathbf{J}_s(0)$.

Sea s_{ri} la cantidad de bloques de Jordan de $\mathbf{J}$ asociados a λ_i de tamaño $r \times r$, con $r = 1, 2, \ldots, p$, siendo p es el tamaño máximo de los bloques de Jordan de la forma $\mathbf{J}_s(\lambda_i)$ presentes en $\mathbf{J}$. Como ya sabemos, para $r = 1, 2, \ldots, p$, s_{ri} también es igual a la cantidad de bloques de Jordan de $\mathbf{J} - \lambda_i \mathbf{I}_n$ de la forma $\mathbf{J}_r(0)$. Notemos que

$$s_{1i} + s_{2i} + \cdots + s_{ri} = b_i.$$

Similarmente, sea s_{ri}' la cantidad de bloques de Jordan de $\mathbf{J}'$ asociados a λ_i de tamaño $r \times r$, con $r = 1, 2, \ldots, p'$, siendo p' es el tamaño máximo de los bloques de Jordan de la forma $\mathbf{J}_s(\lambda_i)$ presentes en $\mathbf{J}'$. También tenemos que

$$s_{1i}' + s_{2i}' + \cdots + s_{ri}' = b_i.$$

En la matriz $(\mathbf{J} - \lambda_i \mathbf{I}_n)^2$, los bloques de la forma $(\mathbf{J}_s(\lambda_j - \lambda_i))^2$ no se anulan, y tiene a $(\lambda_j - \lambda_i)^2$ en su diagonal. Note que por cada bloque de la forma $(\mathbf{J}_r(0))^2$, con $r > 1$, hay dos columnas nulas en $(\mathbf{J} - \lambda_i \mathbf{I}_n)^2$, y cada bloque de la forma $(\mathbf{J}_1(0))^2$ solo aporta una columna nula en $(\mathbf{J} - \lambda_i \mathbf{I}_n)^2$. Luego la cantidad de columnas nulas en $(\mathbf{J} - \lambda_i \mathbf{I}_n)^2$ es

$$s_{1i} + 2(s_{2i} + \cdots + s_{pi}) = s_{1i} + 2(b_i - s_{1i}) = 2b_i - s_{1i}.$$

Ahora, $\mathcal{N}((\mathbf{J} - \lambda_i \mathbf{I}_n)^2)$ es el subespacio propio generalizado de $\mathbf{J}$ de orden 2 asociado a λ_i, y su dimensión es

$$(s_{1i} + s_{2i} + \cdots + s_{ri}) + (s_{2i} + \cdots + s_{ri}) = 2b_i - s_{1i}.$$

La ecuación

$$\dim(\mathcal{N}((\mathbf{J} - \lambda_i \mathbf{I}_n)^2)) + \mathrm{rk}(\mathcal{N}((\mathbf{J} - \lambda_i \mathbf{I})^2)) = n,$$

implica que

$$\mathrm{rk}(\mathcal{N}((\mathbf{J} - \lambda_i \mathbf{I}_n)^2)) = n - 2b_i + s_{1i}.$$

De manera similar,

$$\mathrm{rk}(\mathcal{N}((\mathbf{J}' - \lambda_i \mathbf{I}_n)^2)) = n - 2b_i + s'_{1i}.$$

Como $(\mathbf{J} - \lambda_i \mathbf{I}_n)^2$ y $(\mathbf{J}' - \lambda_i \mathbf{I}_n)^2$ son semejantes, sus rangos son iguales, esto es,

$$n - 2b_i + s_{1i} = n - 2b_i + s'_{1i},$$

de donde $s_{1i} = s'_{1i}$. Así, las matrices $\mathbf{J}$ y $\mathbf{J}'$ tienen la misma cantidad de bloques de Jordan asociados a λ_i de tamaño 1×1.

Ahora, de forma análoga, en la matriz $(\mathbf{J} - \lambda_i \mathbf{I}_n)^3$, los bloques de la forma $(\mathbf{J}_s(\lambda_j - \lambda_i))^3$ no se anulan, y tiene a $(\lambda_j - \lambda_i)^3$ en su diagonal. Por cada bloques de la forma $(\mathbf{J}_r(0))^3$, con $r > 2$ hay tres columnas nulas en $(\mathbf{J} - \lambda_i \mathbf{I}_n)^3$, más las columnas nulas de los bloques de tamaños 1×1 y las columnas nulas aportadas por todos los bloques de tamaño 2×2. Por lo tanto, la cantidad de columnas nulas de $(\mathbf{J} - \lambda_i \mathbf{I}_n)^3$ es

$$s_{1i} + 2s_{2i} + 3(s_{3i} + \cdots + s_{pi}) = 3b_i - 2s_{1i} - s_{2i}.$$

Análogamente a como se hizo antes, de esto se sigue que

$$\mathrm{rk}((\mathbf{J} - \lambda_i \mathbf{I}_n)^3) = n - 3b_i + 2s_{1i} + s_{2i};$$

para la matriz $(\mathbf{J}' - \lambda_i \mathbf{I}_n)^3$ se tiene

$$\mathrm{rk}((\mathbf{J}' - \lambda_i \mathbf{I}_n)^3) = n - 3b_i + 2s'_{1i} + s'_{2i} = n - 3b_i + 2s_{1i} + s'_{2i},$$

y como $(\mathbf{J} - \lambda_i \mathbf{I}_n)^3$ y $(\mathbf{J}' - \lambda_i \mathbf{I}_n)^3$ son semejantes, tienen el mismo rango, es decir

$$n - 3b_i + 2s_{1i} + s_{2i} = n - 3b_i + 2s_{1i} + s'_{2i},$$

y se sigue que $s_{2i} = s'_{2i}$.

Continuando con este proceso inductivo encontramos que $s_{ri} = s'_{ri}$, $r = 1, 2, \ldots, p$; lo que muestra que $p \leq p'$. Pero el mismo argumento intercambiando $\mathbf{J} - \lambda_i \mathbf{I}_n$ con $\mathbf{J}' - \lambda_i \mathbf{I}_n$ lleva a que $p' \leq p$ y por lo tanto $p = p'$.

Llegamos a la conclusión de que en $\mathbf{J} - \lambda_i \mathbf{I}_n$ y en $\mathbf{J}' - \lambda_i \mathbf{I}_n$ hay bloques de los mismos tamaños de la forma $\mathbf{J}_s(0)$ y en las mismas cantidades, lo cual implica que en $\mathbf{J}$ y en $\mathbf{J}'$ existen bloques de los mismos tamaños de la forma $\mathbf{J}_s(\lambda_i)$ y en las mismas cantidades. Lo anterior se tiene para todo $i \in \{1, \ldots, k\}$, así que los bloques de Jordan de $\mathbf{J}'$ se obtienen mediante una permutación de los bloques de Jordan de $\mathbf{J}$, y esto termina la demostración. $\square$

Corolario 4.2.7. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\mathbf{J}, \mathbf{J}' \in \mathbb{C}^{n\times n}$ matrices de Jordan. Entonces $\mathbf{J}$ y $\mathbf{J}'$ están en la misma clase de semejanza de $\mathbf{A}$ si y solo si $\mathbf{J}'$ se obtiene al realizar una permutación de los bloques de Jordan de $\mathbf{J}$.

Demostración. El corolario es consecuencia inmediata de los Teoremas 4.2.5, 4.2.6 y la propiedad transitiva de la semejanza de matrices. $\square$

El Corolario 4.2.7 tiene algunas implicaciones interesantes que mencionamos a continuación:

1. Sea $\mathbf{J} \in \mathbb{C}^{n\times n}$ una matriz de Jordan con k bloques de Jordan y sean
$$\mathbf{J}_{n_1}(\lambda_1), \ldots, \mathbf{J}_{n_s}(\lambda_s)$$
los bloques de Jordan sobre la diagonal principal de $\mathbf{J}$ que son distintos (si $\lambda_i = \lambda_j$, necesariamente $n_i \neq n_j$). Digamos que cada bloque de estos se repite $t_1, \ldots, t_s$ veces, respectivamente. Entonces la clase de semejanza de $\mathbf{J}$ tiene exactamente
$$\frac{k!}{t_1! \cdots t_s!}$$
matrices de Jordan distintas.

2. Si en una clase de semejanza hay una matriz diagonal, entonces en dicha clase todas las matrices de Jordan son diagonales.

3. Si una matriz de Jordan no es diagonal, entonces no es diagonalizable.

Corolario 4.2.8. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ y $\mathbf{J}$ una matriz de Jordan semejante a $\mathbf{A}$. Entonces los bloques de Jordan de mayor tamaño en $\mathbf{J}$ asociados a λ son de tamaño $i_{\mathbf{A}}(\lambda) \times i_{\mathbf{A}}(\lambda)$.

Demostración. Como $\mathbf{A}$ y $\mathbf{J}$ son semejantes, entonces existe $\mathbf{P} \in \mathbb{C}^{n\times n}$ invertible tal que

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}.$$

Por el Corolario 4.2.7, podemos asumir que los bloque de Jordan en $\mathbf{J}$ asociados a λ están ubicados consecutivamente sobre parte más alta de la diagonal principal de $\mathbf{J}$. Conectando adecuadamente ideas presentes en las demostraciones del Teorema 3.2.3 y del Teorema de Jordan, y por lo que hemos supuesto, solo en la parte inicial de $\mathbf{P}$ hay cadenas de Jordan asociadas a λ y están consecutivas. Además, cada cadena determina un bloque de Jordan en $\mathbf{J}$ asociados a λ y de tamaño igual a la longitud de la cadena correspondiente. Como hay cadenas de longitud $i_{\mathbf{A}}(\lambda)$ y no hay cadenas de mayor longitud, entonces los bloques de Jordan de mayor tamaño en $\mathbf{J}$ asociados a λ son de tamaño $i_{\mathbf{A}}(\lambda) \times i_{\mathbf{A}}(\lambda)$. □

Los siguientes dos resultados son consecuencia del Teorema 3.2.3 y su demostración.

Teorema 4.2.9. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lambda \in \operatorname{Spec}(\mathbf{A})$ con $i_{\mathbf{A}}(\lambda) = p$,

$$\dim(\mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_n)^p) - \dim(\mathcal{N}(\mathbf{A} - \lambda\mathbf{I}_n)^{p-1}) = k$$

y $\mathbf{J}$ una matriz de Jordan semejante a $\mathbf{A}$. Entonces $\mathbf{J}$ tiene exactamente k bloques de Jordan de tamaño $p \times p$ asociados a λ.

Demostración. Como $\dim(\mathcal{N}(\mathbf{A}-\lambda\mathbf{I}_n)^p)-\dim(\mathcal{N}(\mathbf{A}-\lambda\mathbf{I}_n)^{p-1}) = k$, entonces existen $\mathbf{x}_1, \ldots, \mathbf{x}_k$ vectores propios generalizados de $\mathbf{A}$ de orden p asociados a λ tales que $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ es linealmente independiente. Entonces, así como en la demostración del Teorema 3.2.3 y en especial en el diagrama (3.3), los vectores $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ generan las cadenas de Jordan de mayor longitud (las primeras k filas), es decir, de longitud p. Las demás cadenas de Jordan en el diagrama tienen longitud menor que p, así que la matriz de Jordan que resulta de este proceso tiene exactamente k bloques de Jordan de tamaño $p \times p$. □

Teorema 4.2.10. Sean $\mathbf{A} \in \mathbb{C}^{n\times n}$, λ un valor propio de $\mathbf{A}$ de índice p, $\mathbf{J}$ una matriz de Jordan semejante a $\mathbf{A}$ y $r \in \{1, \ldots, p-1\}$. Entonces $\mathbf{J}$ tiene bloques de Jordan de tamaño $r \times r$ asociados a λ si

$$\mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}) > \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r+1}) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r),$$

y $\mathbf{J}$ no tiene bloques de Jordan de tamaño $r \times r$ asociados a λ si

$$\mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}) = \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r+1}) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r).$$

Cuando

$$\mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}) > \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r+1}) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r),$$

$\mathbf{J}$ tiene exactamente

$$2\mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^r) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r-1}) - \mathrm{rk}((\mathbf{A}-\lambda\mathbf{I}_n)^{r+1})$$

bloques de Jordan de tamaño $r \times r$ asociados a λ.

Demostración. Podemos asumir que $\mathbf{J}$ es la matriz que se obtiene mediante las construcciones de los Teoremas 3.2.3 y 4.1.2. Usaremos la notación usada en la demostración del Teorema 3.2.3. A continuación reproducimos el diagrama (3.3). Para este diagrama, $t = t_p + t_{p-1} + \cdots + t_1$:

$$\begin{array}{ccccc}
\mathbf{x}_1^{(1)}, & \mathbf{x}_2^{(1)}, & \ldots, & \mathbf{x}_{p-1}^{(1)}, & \mathbf{x}_p^{(1)}, \\
\vdots & \vdots & & \vdots & \vdots \\
\mathbf{x}_1^{(t_p)}, & \mathbf{x}_2^{(t_p)}, & \ldots, & \mathbf{x}_{p-1}^{(t_p)}, & \mathbf{x}_p^{(t_p)}, \\
\mathbf{x}_1^{(t_p+1)}, & \mathbf{x}_2^{(t_p+1)}, & \ldots, & \mathbf{x}_{p-1}^{(t_p+1)}, & \\
\vdots & \vdots & & \vdots & \\
\mathbf{x}_1^{(t_p+t_{p-1})}, & \mathbf{x}_2^{(t_p+t_{p-1})}, & \cdots, & \mathbf{x}_{p-1}^{t_p+t_{p-1}}, & \\
\vdots & \vdots & & & \\
\mathbf{x}_1^{(t-t_1)}, & \mathbf{x}_2^{(t-t_1)}, & & & \\
\mathbf{x}_1^{(t-t_1+1)}, & & & & \\
\vdots & & & & \\
\mathbf{x}_1^{(t)}, & & & &
\end{array} \tag{4.2}$$

Recordemos que $\mathcal{U}_i = \mathcal{N}((\mathbf{A}-\lambda\mathbf{I}_n)^i)$, para $i = 1, 2, \ldots, p$. Ahora, sea $r \in \{1, \ldots, p-1\}$.

Como se hizo en la construcción en la demostración del Teorema 3.2.3, la $j-$ésima columna de izquierda a derecha en (4.2) está formada por vectores propios generalizados de orden j, donde $j \in \{1, \dots, p\}$. Cada fila del diagrama (4.2) es una cadena de Jordan; las primeras t_p filas de arriba son cadenas de Jordan de longitud p asociadas a λ. Las t_{p-1} filas que siguen son cadenas de Jordan de longitud $p-1$; las t_{p-2} siguientes son cadenas de Jordan de longitud $p-2$, y así sucesivamente.

Así, la cantidad de cadenas de Jordan presentes en (4.2) de longitud r es precisamente t_r, donde $r \in \{1, 2, \dots, p\}$. Cada cadena de Jordan en (4.2) de longitud r corresponde a un bloque de Jordan de $\mathbf{J}$ de tamaño $r \times r$, así que la cantidad de bloques de Jordan en $\mathbf{J}$ asociados a λ de tamaño $r \times r$ es t_r, $r = 1, \dots, p$.

Sea $r \in \{1, \dots, p-1\}$. Entonces tenemos que

$$\mathcal{U}_{r-1} \oplus \text{gen}\{\mathbf{x}_r^{(1)}, \dots, \mathbf{x}_r^{(t_p+\cdots+t_{r+1})}, \mathbf{x}_r^{(t_p+\cdots+t_{r+1}+1)}, \dots, \mathbf{x}_r^{(t_p+\cdots+t_{r+1}+t_r)}\} = \mathcal{U}_r,$$

así que se sigue que

$$\dim(\mathcal{U}_{r-1}) + \dim(\text{gen}\{\mathbf{x}_r^{(1)}, \dots \mathbf{x}_r^{(t_p+\cdots+t_{r+1}+t_r)}\}) = \dim(\mathcal{U}_r),$$

y como

$$\dim(\text{gen}\{\mathbf{x}_r^{(1)}, \dots \mathbf{x}_r^{(t_p+\cdots+t_{r+1}+t_r)}\}) = t_p + \cdots + t_{r+1} + t_r$$

obtenemos que

$$\dim(\mathcal{U}_{r-1}) + (t_p + \cdots + t_{r+1} + t_r) = \dim(\mathcal{U}_r). \tag{4.3}$$

Por otro lado

$$\mathcal{U}_r \oplus \text{gen}\{\mathbf{x}_{r+1}^{(1)}, \dots, \mathbf{x}_{r+1}^{(t_p+\cdots+t_{r+1})}\} = \mathcal{U}_{r+1},$$

de donde

$$\dim(\mathcal{U}_r) + \dim(\text{gen}\{\mathbf{x}_{r+1}^{(1)}, \dots, \mathbf{x}_{r+1}^{(t_p+\cdots+t_{r+1})}\}) = \dim(\mathcal{U}_{r+1})$$

y como $\dim(\text{gen}\{\mathbf{x}_{r+1}^{(1)}, \dots, \mathbf{x}_{r+1}^{(t_p+\cdots+t_{r+1})}\}) = t_p + \cdots + t_{r+1}$, resulta que

$$t_p + \cdots + t_{r+1} = \dim(\mathcal{U}_{r+1}) - \dim(\mathcal{U}_r).$$

Sustituyendo esto último en (4.3) obtenemos que

$$\dim(\mathcal{U}_{r-1}) + \dim(\mathcal{U}_{r+1}) - \dim(\mathcal{U}_r) + t_r = \dim(\mathcal{U}_r).$$

y esto muestra que

$$t_r = 2\dim(\mathcal{U}_r) - \dim(\mathcal{U}_{r-1}) - \dim(\mathcal{U}_{r+1}).$$

Finalmente, $t_r > 0$ es equivalente a que

$$\dim(\mathcal{U}_r) - \dim(\mathcal{U}_{r-1}) > \dim(\mathcal{U}_{r+1}) - \dim(\mathcal{U}_r)$$

y si esta condición se satisface, entonces existen exactamente

$$t_r = 2\dim(\mathcal{U}_r) - \dim(\mathcal{U}_{r-1}) - \dim(\mathcal{U}_{r+1})$$

bloques de Jordan de tamaño $r \times r$. $\square$

Bibliografía

[1] J. Aldrich, *Eigenvalue, eigenfunction, eigenvector, and related terms*, Jeff Miller (Editor), Earliest Known Uses of Some of the Words of Mathematics, 2006.

[2] F. Brechenmacher, A Controversy and the writing of a history - The discusión of "small oscilaciones"(1760-1860) from the standpoint of the Controversy between Jordan and Kronecker (1874), *Bulletin of the Belgian Mathematical Society,* (*Francia*), **1**, No. 13, (2006), 1-4.

[3] F. Brechenmacher, *Histoire du théorème de Jordan de la décomposition matricielle (1870-1930). Formes de représentation et méthodes de décomposition,* Ecole des Hautes Etudes en Sciences Sociales (EHESS), Francia, 2006.

[4] J. Fraleigh, *Álgebra Abstracta*, Addison-Wesley, Wilmington, 1988.

[5] S. Grossman, *Álgebra lineal*, McGraw, México, 2012.

[6] T. Hawkins, Cauchy and the spectral theory of matrices, *Historia Mathematica*, **2**, No. 1, (1975), 1-29.

[7] I.N Herstein, *Álgebra moderna*, Editorial Trillas, México, 1980.

[8] K. Hoffman, R. Kunze, *Álgebra lineal*, Prentice Hall, México, 1973.

[9] R. Horn, C. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[10] P. Lancaster, M. Tismenetsky, *The theory of matrices*, Academic Press, San Diego, 1985.

[11] C. Meyer, *Matrix Analyisis and Applied Linear Algebra*, SIAM, New York, 2010.